HUMAN BIOLOGY
a record of research

MAY, 1951

VOL. 23 No. 2

DATA ON THE GROWTH OF CHILDREN DURING THE FIRST YEAR AFTER BIRTH

BY HELEN THOMPSON

The Clinic of Child Development, Yale University

AS PART of the study on infancy begun in 1927 at the Child Development Clinic of Yale University, data on physical growth were obtained. Measurements of nine dimensions were made every four weeks on healthy infants from 6 weeks to 52 weeks old. The series is a mixed longitudinal one; the majority of children appear more than once in the records, but only a few continue uninterruptedly from 8 to 52 weeks.

These measurements have already been reported cross-sectionally (1); that is, the overall mean and standard deviation for each age period was given, to serve as standards for size of other infants in similar or different circumstances. Changes in body proportions during this first year have also been discussed (2). The specifically longitudinal aspect of the data, however, has not yet been taken advantage of, and the actual raw figures have never been published. This paper repairs both these omissions.

The data in this study were collected in connection with the Yale University Clinic of Child Development Normative Study of Infancy, Dr. Arnold Gesell, Director, with whose permission this article is

published. The analysis was suggested by Dr. J. M. Tanner and the statistics compiled by Dr. Tanner and his assistant, Mr. R. H. Whitehouse, to whom the author expresses her appreciation.

In Tables 1 and 2 are given the means and standard deviations of the increments of growth over four-week periods. These figures are calculated from the actual individual increments; they are not the same as those obtained by subtracting one cross-sectional mean from the next one. This point is elaborated and discussed in the accompanying paper by Tanner on the reporting of growth data (3). The number of children for each incremental period is distressingly small, but better, after all, than nothing; and no incremental data of this sort except for length and weight have yet appeared. The work-up of this sort of data should normally include the correlations between all the increments, but here the numbers are much too small to warrant working them out. Likewise no individual curves have been fitted, as very few children have long-enough records for satisfactory estimates either of form or constants to be made. It should also be noted that ideally for mean increments such as these to be valid standards, the individual monthly increments should be evenly distributed over the months in the year; i. e. the 8 to 12 week increment should have some January increments, some February and so on, since increments vary with the season. The present data are too few for the seasonal effect to be taken into account.

The remaining tables (Tables 3 and 4) give the raw data; no others appear to have been published for this age with the exception of Davenport's (4), which are slightly smoothed and thus depart somewhat from the raw state. In these tables a few figures which seem to be clearly erroneously recorded are enclosed in parentheses: these were deleted in the calculations for Tables 1 and 2. Both the material and the methods of measurement have been fully set forth previously (1). Only the salient points are recapitulated here.

TABLE 1

MATERIAL

The records for 32 male and 36 female infants are sufficiently extensive to warrant recording; the number of attendances for these children varies from 3 to 24. They were from homes of average socio-economic status in the New Haven area. All the parents were American-born, the majority of the grandparents also. The infants were all born at full term, were healthy, and developed normally in their behavior.

TABLE 1

Increments of growth during the first year of life: boys

MEASUREMENT		AGE IN WEEKS 8-12	12-16	16-20	20-24	24-28	28-32	32-36	36-40	40-44	44-48	48-52
Vertex height (soles—vertex) (cm)	N	11	11	10	11	9	9	9	12	15	14	14
	Mn	2.98	2.41	2.05	1.93	1.02	1.60	0.77	1.32	1.11	1.46	1.40
	σ	1.18	1.14	0.44	0.63	0.24	0.60	0.58	0.51	0.54	0.59	0.44
Suprasternal height (soles—suprasternal notch)	N	11	12	10	11	9	9	9	12	15	14	14
	Mn	2.50	1.58	1.51	1.24	1.02	1.18	0.63	0.75	1.11	1.29	1.19
	σ	1.29	1.00	0.64	0.67	0.34	0.51	0.64	0.64	0.54	0.67	0.78
Symphysis height (soles—symphysis)	N	12	11	10	11	9	9	9	12	15	14	14
	Mn	1.41	1.41	0.93	0.80	0.69	1.09	0.76	0.49	0.68	0.72	0.64
	σ	0.54	0.86	0.34	0.61	0.45	0.68	0.86	0.53	0.58	0.64	0.38
Vertex—suprasternal notch	N	10	11	10	11	9	9	9	11	15	14	14
	Mn	0.58	0.90	0.54	0.69	0.00	0.42	0.14	0.51	0.16	0.01	0.27
	σ	0.78	0.63	0.68	0.49	0.37	0.38	0.33	0.39	0.62	0.64	0.80
Vertex—symphysis	N	11	10	10	11	9	9	9	12	15	14	13
	Mn	1.64	0.92	1.12	1.13	0.33	0.51	0.02	0.83	0.43	0.74	0.74
	σ	1.08	0.66	0.48	0.58	0.51	0.88	0.76	0.62	0.75	0.93	0.65
Suprasternal notch—symphysis	N	10	10	10	11	9	9	9	11	15	14	14
	Mn	1.16	−0.09	0.58	0.44	0.34	0.09	−0.12	0.38	0.25	0.71	0.46
	σ	1.24	0.74	0.55	0.58	0.54	0.74	0.70	0.80	0.91	0.91	0.67
Biacromial diameter	N	12	12	9	10	7	7	8	10	15	14	13
	Mn	0.65	0.71	0.87	0.21	0.37	0.21	0.30	0.41	0.49	0.41	0.36
	σ	0.74	1.01	0.52	0.58	0.45	0.24	0.40	0.67	0.45	0.41	0.32
Chest breadth	N	10	11	10	11	7	7	9	10	13	13	13
	Mn	1.01	0.45	0.51	0.28	0.49	0.59	0.21	0.32	0.22	0.28	0.34
	σ	0.63	0.66	0.46	0.44	0.12	0.37	0.29	0.29	0.25	0.36	0.28
Chest circumference	N	12	11	9	10	7	7	7	9	13	12	13
	Mn	1.70	1.53	1.19	1.52	0.64	0.56	0.21	2.22	0.62	0.37	0.47
	σ	1.07	1.72	1.03	0.94	0.79	0.76	0.50	1.12	1.47	0.78	0.88
Bicristal diameter	N	12	12	10	11	9	9	9	10	15	14	14
	Mn	0.73	0.27	0.38	0.25	0.11	0.28	0.16	0.20	0.21	0.28	0.12
	σ	0.65	0.36	0.36	0.25	0.57	0.52	0.21	0.24	0.31	0.32	0.25
Head circumference	N	11	11	10	11	8	8	9	11	16	14	13
	Mn	1.32	1.11	0.96	1.03	0.35	0.90	0.43	0.48	0.48	0.51	0.25
	σ	0.41	0.38	0.30	0.26	0.22	0.32	0.21	0.26	0.21	0.20	0.21
Body weight (kg)	N	11	11	10	11	9	9	9	12	15	14	14
	Mn	0.77	0.71	0.57	0.72	0.36	0.47	0.39	0.43	0.49	0.22	0.36
	σ	0.19	0.25	0.44	0.28	0.20	0.24	0.14	0.25	0.35	0.22	0.20

TABLE 2

Increments of growth during the first year of life: girls

MEASUREMENT		AGE IN WEEKS 8-12	12-16	16-20	20-24	24-28	28-32	32-36	36-40	40-44	44-48	48-52
Vertex height (soles—vertex) (cm)	N	11	13	8	11	15	10	11	14	12	15	15
	Mn	2.26	2.46	2.30	1.95	1.26	1.76	1.26	1.47	1.03	1.22	1.10
	σ	0.73	0.86	0.49	0.65	0.92	0.68	0.65	0.48	0.58	0.62	0.48
Suprasternal height (soles—suprasternal notch)	N	10	11	8	11	15	10	11	14	12	14	14
	Mn	2.06	1.63	1.95	1.58	0.95	1.25	0.90	1.16	0.73	1.11	1.04
	σ	1.02	.058	0.48	0.60	1.03	0.61	0.83	0.77	0.57	0.78	0.77
Symphysis height (soles—symphysis)	N	10	12	7	12	14	9	11	13	12	15	15
	Mn	1.29	1.08	1.16	0.85	1.00	0.68	1.02	0.62	0.74	0.60	0.58
	σ	0.37	0.32	0.61	0.37	0.67	0.45	0.81	0.52	0.46	0.42	0.37
Vertex—suprasternal notch	N	10	11	8	11	15	10	11	14	12	14	14
	Mn	0.22	0.85	0.35	0.36	0.36	0.43	0.36	0.31	0.30	0.11	0.09
	σ	0.74	0.61	0.42	0.56	0.54	0.35	0.78	0.51	0.47	0.48	0.84
Vertex—symphysis	N	10	12	8	11	14	9	11	13	12	15	15
	Mn	1.19	1.36	1.19	1.31	0.41	0.90	0.25	0.89	0.34	0.53	0.64
	σ	0.84	1.09	0.81	0.59	1.10	0.83	1.16	0.65	0.48	0.55	0.49
Suprasternal notch—symphysis	N	10	11	8	11	14	9	11	13	12	14	14
	Mn	0.77	0.48	0.74	0.80	0.06	0.49	−0.13	0.58	−0.02	0.51	0.66
	σ	0.98	0.82	0.72	0.71	1.11	0.56	1.42	0.97	0.60	0.77	0.96
Biacromial diameter	N	11	13	8	10	13	9	10	12	10	14	15
	Mn	0.35	0.78	0.58	0.29	0.32	0.37	0.55	0.35	0.39	0.51	0.30
	σ	0.70	0.90	0.46	0.51	0.47	0.37	0.44	0.28	0.28	0.54	0.42
Chest breadth	N	11	13	7	10	14	8	10	11	11	15	14
	Mn	0.55	0.80	0.41	0.41	0.34	0.44	0.36	0.30	0.27	0.28	0.19
	σ	0.71	0.95	0.48	0.43	0.47	0.29	0.27	0.30	0.33	0.39	0.37
Chest circumference	N	10	12	8	9	14	9	9	13	12	15	15
	Mn	1.32	2.16	1.58	0.84	1.17	0.66	1.28	0.55	0.64	0.88	0.73
	σ	1.10	0.87	1.16	0.62	0.74	1.13	1.38	0.91	0.93	0.52	0.70
Bicristal diameter	N	11	13	7	11	15	10	11	14	12	15	15
	Mn	0.45	0.46	0.27	0.34	0.14	0.42	0.25	0.28	0.14	0.23	0.07
	σ	0.37	0.36	0.20	0.70	0.37	0.53	0.19	0.30	0.16	0.43	0.32
Head circumference	N	11	12	8	11	15	10	11	14	12	15	14
	Mn	1.05	1.13	1.06	0.77	0.77	0.60	0.65	0.41	0.34	0.39	0.39
	σ	0.36	0.57	0.20	0.24	0.38	0.30	0.24	0.26	0.36	0.34	0.31
Body weight (kg)	N	13	13	8	10	14	10	11	14	12	15	15
	Mn	0.68	0.70	0.64	0.56	0.44	0.47	0.32	0.36	0.24	0.44	0.20
	σ	0.19	0.34	0.21	0.23	0.23	0.21	0.18	0.21	0.23	0.29	0.30

TABLE 3

Raw data on males

AGE	VERTEX HEIGHT	SUPRA-STERNAL HEIGHT	SYM-PHYSIS HEIGHT	BIA-CROMIAL DIAM-ETER	CHEST BREADTH	CHEST CIRCUM.	BI-CRISTAL DIAM-ETER	HEAD CIRCUM.	WEIGHT
Subject M.1									
8 wks.	60.3	44.2	23.5	13.6	11.7	38.1	10.2	40.6	5.54
12 wks.	62.6	46.4	25.1	14.8	12.9	40.2	10.9	41.6	6.17
16 wks.	64.6	47.8	25.6	14.6	13.1	41.1	10.9	42.8	6.72
20 wks.	67.1	49.4	26.7	14.8	13.3	41.8	11.5	43.5	7.22
24 wks.	68.4	50.4	27.6	15.4	13.9	42.4	11.5	44.5	7.58
28 wks.	68.8	51.2	28.4	15.8	14.2	42.4	11.6	44.8	7.80
32 wks.	70.9	52.5	29.7	15.9	14.4	43.8	11.7	45.5	8.29
36 wks.	71.9	53.0	30.1	16.4	14.5	44.2	12.2	46.0	8.86
40 wks.	74.0	54.8	30.3	16.8	15.3	47.4	12.6	46.6	9.63
44 wks.	75.2	55.9	30.8	17.4	15.5	48.0	13.2	47.1	10.90
48 wks.	77.3	58.3	31.7	18.1	15.9	49.2	13.4	47.5	11.52
52 wks.	78.3	59.3	32.8	18.7	16.3	51.0	13.7	48.0	12.01
56 wks.	80.5	60.5	33.9	19.6	16.6	51.4	13.7	48.1	12.64
80 wks.	86.2	65.3	37.8	20.8	—	51.6	14.4	49.8	14.16
2 yrs.	92.6	70.5	42.3	20.4	18.6	52.2	15.8	50.8	14.78
3 yrs.	103.2	79.8	48.8	22.1	18.6	53.8	17.3	51.4	17.74
5 yrs.	118.9	93.7	—	25.5	18.9	—	19.4	52.5	23.04
6 yrs.	126.0	100.2	63.9	26.2	19.7	60.2	20.1	53.2	25.51
7 yrs.	131.9	105.8	67.5	28.2	21.0	62.4	21.2	52.9	—
8 yrs.	139.2	112.7	72.7	29.6	20.9	65.9	22.2	53.5	32.69
10 yrs.	151.7	124.4	—	35.4	23.3	69.8	23.4	54.2	40.46
Subject M.2									
6 wks.	55.1	—	20.7	13.2	9.2	34.0	11.2	38.7	4.60
8 wks.	56.2	41.6	21.6	12.6	9.9	35.6	11.2	39.5	5.00
12 wks.	—	43.7	23.7	13.2	12.2	38.8	10.7	41.2	5.82
16 wks.	62.5	46.1	24.9	14.8	11.9	38.2	10.4	42.3	6.05
20 wks.	64.3	47.4	26.2	15.5	12.2	38.8	10.9	42.9	6.52
24 wks.	66.0	48.7	26.8	15.6	12.7	40.4	11.4	43.8	6.91
28 wks.	67.0	49.6	27.2	15.7	13.1	40.7	11.6	44.1	7.18
32 wks.	69.8	51.7	28.1	16.2	13.5	41.0	11.9	45.5	7.65
36 wks.	69.1	51.1	28.2	15.9	13.8	41.9	11.9	45.7	7.98
40 wks.	71.0	53.0	28.7	16.8	14.2	43.5	12.1	46.0	8.28
44 wks.	72.6	53.8	30.3	16.7	14.2	44.0	12.2	46.6	8.48
48 wks.	73.0	54.2	30.9	16.9	14.3	44.1	12.3	46.9	8.60
52 wks.	74.1	55.1	31.9	17.3	14.9	44.9	12.6	47.2	8.99
56 wks.	75.2	55.7	32.7	17.2	15.1	45.5	12.8	47.8	9.30
60 wks.	76.6	57.0	33.2	17.7	15.4	45.3	13.2	48.0	9.50
80 wks.	80.6	61.2	36.2	18.7	15.8	48.7	13.9	48.9	10.12
3 yrs.	93.2	71.9	43.6	20.1	—	—	16.1	50.2	12.40
4 yrs.	99.5	77.6	47.6	21.6	16.0	49.2	16.8	50.7	13.30
5 yrs.	104.1	82.1	51.1	21.7	16.8	49.6	16.4	51.1	13.33
6 yrs.	108.5	85.7	53.0	22.8	17.0	52.8	17.9	50.9	15.13
7 yrs.	112.5	89.7	56.2	23.8	17.5	53.3	18.4	50.8	14.77
8 yrs.	116.0	92.5	58.1	24.5	18.7	54.5	18.9	51.2	15.90
9 yrs.	119.2	94.9	60.0	24.3	18.3	55.1	19.2	51.1	16.71
10 yrs.	122.0	98.9	—	24.4	—	—	19.5	51.7	18.21

TABLE 3—Continued

Raw data on males

AGE	VERTEX HEIGHT	SUPRA-STERNAL HEIGHT	SYM-PHYSIS HEIGHT	BIA-CROMIAL DIAM-ETER	CHEST BREADTH	CHEST CIRCUM.	BI-CRISTAL DIAM-ETER	HEAD CIRCUM.	WEIGHT
Subject M.3									
8 wks.	57.4	42.0	22.5	12.4	—	37.6	9.5	37.6	5.10
12 wks.	63.0	48.0	24.5	13.2	11.2	38.8	11.8	39.5	5.92
16 wks.	62.9	47.3	24.6	15.8	13.4	44.5	11.6	40.6	6.87
20 wks.	64.8	48.3	25.0	16.6	13.6	44.5	11.7	41.8	7.47
24 wks.	67.2	50.4	26.3	16.1	13.7	46.7	11.8	43.0	8.29
28 wks.	68.2	51.4	26.7	16.6	14.2	46.9	11.9	43.2	8.46
32 wks.	69.5	52.4	28.7	16.9	14.8	47.3	12.3	44.4	8.92
36 wks.	70.4	52.7	28.3	16.6	14.5	47.2	12.2	44.5	9.33
40 wks.	71.7	53.4	29.7	16.8	14.8	47.9	12.2	44.9	9.61
44 wks.	72.0	54.2	31.4	17.2	14.8	47.4	12.3	45.3	9.80
48 wks.	73.9	55.4	31.3	17.5	15.7	47.6	12.7	45.9	10.10
56 wks.	76.7	57.7	32.1	18.1	15.8	49.1	13.1	46.5	10.60
80 wks.	81.8	62.8	35.7	17.8	16.7	50.7	13.6	47.4	11.84
2 yrs.	83.4	65.3	39.9	19.7	16.8	51.8	14.8	48.5	12.77
3 yrs.	93.6	71.5	41.4	20.3	—	—	16.6	48.4	15.52
4 yrs.	102.2	79.4	48.0	22.1	18.5	56.2	17.0	49.8	17.50
5 yrs.	108.1	84.5	52.1	23.7	18.8	56.7	17.6	50.0	18.94
6 yrs.	115.1	90.6	56.3	24.9	—	58.5	18.6	50.9	20.46
7 yrs.	121.7	96.7	60.7	26.0	—	60.5	19.6	50.8	23.13
8 yrs.	128.3	102.7	63.9	26.2	—	61.8	20.4	51.4	26.79
10 yrs.	138.0	112.4	70.8	27.0	—	65.4	21.6	52.0	31.10
Subject M.4									
16 wks.	62.0	46.5	24.0	13.2	11.1	40.4	9.6	41.7	5.82
20 wks.	63.5	46.7	24.2	14.5	12.1	42.1	10.2	42.4	6.34
24 wks.	64.9	48.0	25.0	15.1	12.1	42.8	10.3	43.2	7.15
28 wks.	66.0	49.6	26.3	—	12.5	42.3	10.6	43.9	7.40
40 wks.	69.3	51.9	27.9	15.6	—	45.1	11.1	45.1	8.46
44 wks.	69.9	52.3	27.9	16.5	13.8	47.4	11.6	45.5	9.14
48 wks.	71.7	54.5	29.7	17.7	14.5	47.9	11.8	46.3	9.69
52 wks.	73.6	55.3	29.9	17.7	15.2	47.8	12.4	46.6	10.21
56 wks.	74.4	56.1	30.0	17.8	15.2	48.6	12.4	47.0	10.46
Subject M.5									
16 wks.	65.6	48.9	27.6	14.4	11.6	42.6	10.9	41.8	7.37
20 wks.	68.2	51.2	28.5	14.7	12.6	44.6	10.5	43.2	8.17
24 wks.	71.2	52.9	29.8	15.4	12.3	47.2	10.7	44.0	8.99
40 wks.	77.1	57.2	33.7	16.6	14.6	46.0	11.8	46.3	10.85
44 wks.	77.9	58.9	33.4	17.7	14.4	50.7	12.0	46.8	11.66
48 wks.	79.6	59.4	35.0	17.6	14.5	51.0	12.4	47.5	11.73
52 wks.	81.0	61.0	35.6	18.1	14.8	49.6	12.1	48.0	12.15
56 wks.	81.9	62.0	36.3	17.8	15.9	51.6	12.9	48.3	12.72
Subject M.6									
16 wks.	64.3	48.2	25.1	14.4	11.6	41.4	9.8	40.2	6.63
20 wks.	66.1	49.6	26.2	15.1	12.6	41.3	10.2	41.0	7.14
24 wks.	68.0	50.3	27.1	15.3	12.6	44.2	10.6	41.7	7.84
36 wks.	72.6	53.9	31.1	16.4	12.7	43.6	11.6	43.7	8.49
40 wks.	72.7	54.2	30.5	16.7	13.1	45.6	11.6	44.1	8.93

TABLE 3—Continued

Raw data on males

AGE	VERTEX HEIGHT	SUPRA-STERNAL HEIGHT	SYM-PHYSIS HEIGHT	BIA-CROMIAL DIAM-ETER	CHEST BREADTH	CHEST CIRCUM.	BI-CRISTAL DIAM-ETER	HEAD CIRCUM.	WEIGHT
44 wks.	74.3	54.6	32.0	17.1	13.9	—	11.8	44.7	9.48
48 wks.	76.4	56.6	32.2	17.3	14.1	45.3	12.5	45.2	9.62
52 wks.	76.8	57.2	33.1	17.9	14.9	46.2	12.6	45.4	9.54
56 wks.	77.7	57.7	33.5	18.3	15.5	45.7	12.6	45.7	9.42
Subject M.7									
16 wks.	61.1	45.8	23.9	13.1	12.1	39.5	10.2	41.1	5.64
20 wks.	64.0	47.0	24.9	14.9	12.6	40.9	10.4	42.8	6.41
36 wks.	71.1	52.8	29.6	—	—	—	—	45.6	8.15
40 wks.	72.6	53.5	29.9	16.1	14.1	43.9	14.1	45.6	8.56
44 wks.	74.2	55.2	30.4	16.2	14.4	44.5	14.4	46.7	8.82
48 wks.	75.4	56.1	30.9	16.8	14.7	46.5	14.7	47.4	9.14
52 wks.	76.8	57.6	31.2	17.6	14.9	46.7	14.9	47.2	9.85
Subject M.8									
16 wks.	64.7	47.9	25.9	15.1	12.3	41.1	10.9	41.8	6.72
32 wks.	71.3	52.6	28.6	15.9	14.1	—	11.7	44.9	9.36
36 wks.	72.9	54.6	30.9	16.6	14.4	44.9	11.7	45.4	9.53
44 wks.	75.7	56.9	33.7	17.8	15.3	45.9	12.2	46.4	10.30
48 wks.	77.6	58.4	33.0	18.1	15.4	46.6	12.3	47.1	10.33
52 wks.	78.6	58.9	33.6	18.2	15.3	46.8	12.3	47.1	10.62
Subject M.9									
16 wks.	62.1	46.3	25.0	14.4	12.5	42.5	9.8	42.5	6.75
20 wks.	64.0	48.7	26.3	(13.6)	12.0	—	10.8	43.2	6.19
24 wks.	65.3	48.5	26.4	14.6	12.5	43.4	11.2	44.5	7.38
28 wks.	66.4	49.1	27.9	15.4	(13.3)	44.7	10.2	45.2	7.74
32 wks.	67.5	50.0	27.7	15.7	12.4	—	10.5	45.6	7.87
48 wks.	72.0	55.0	31.1	16.1	—	45.7	11.5	47.3	8.92
52 wks.	73.7	55.5	31.5	17.1	12.3	47.1	11.6	47.6	9.52
Subject M.10									
28 wks.	68.5	50.7	27.2	16.1	13.4	45.8	11.7	45.8	9.29
40 wks.	70.8	52.7	29.1	15.6	13.7	46.6	12.6	47.1	9.85
44 wks.	72.0	53.5	29.4	16.6	13.6	45.5	12.4	47.7	10.13
52 wks.	74.8	55.8	31.5	15.4	14.1	47.4	12.3	48.5	10.66
Subject M.11									
16 wks.	63.2	45.9	24.2	13.6	11.1	41.1	10.4	42.2	6.92
20 wks.	64.8	48.1	24.8	15.1	12.0	44.5	11.0	43.0	7.82
24 wks.	67.6	50.0	25.4	15.3	13.3	44.2	11.1	44.3	8.79
40 wks.	73.8	54.9	29.7	16.6	13.1	47.9	12.6	47.0	11.09
48 wks.	76.4	57.2	30.7	17.5	14.1	48.3	13.0	48.2	12.14
Subject M.12									
20 wks.	65.4	48.6	26.5	15.2	12.4	43.6	11.6	43.4	7.16
24 wks.	67.6	50.2	28.6	14.7	12.6	45.3	12.2	44.9	8.17
28 wks.	68.6	51.0	28.7	(16.7)	—	—	13.5	45.2	8.58
32 wks.	70.2	51.5	28.8	15.1	13.8	46.6	12.6	45.8	9.06
52 wks.	75.6	56.0	33.0	15.4	14.1	47.0	12.9	47.6	9.87

TABLE 3—Continued

Raw data on males

AGE	VERTEX HEIGHT	SUPRA-STERNAL HEIGHT	SYM-PHYSIS HEIGHT	BIA-CROMIAL DIAM-ETER	CHEST BREADTH	CHEST CIRCUM.	BI-CRISTAL DIAM-ETER	HEAD CIRCUM.	WEIGHT
Subject M.13									
20 wks.	66.5	49.6	27.1	15.7	13.6	44.5	11.6	42.5	8.25
24 wks.	67.9	50.0	26.8	15.1	13.3	45.6	12.2	43.2	8.53
28 wks.	69.0	51.5	27.6	15.4	13.9	—	11.9	43.2	8.70
36 wks.	71.6	53.8	28.8	—	—	—	—	—	9.41
40 wks.	72.5	53.5	29.6	17.4	14.2	46.4	12.6	45.5	9.25
44 wks.	73.7	55.9	30.4	17.1	14.6	46.9	12.5	45.7	10.22
48 wks.	75.6	56.7	31.6	18.1	14.2	46.3	12.8	46.5	10.00
52 wks.	77.2	58.0	32.2	17.9	14.6	46.6	12.4	46.9	10.44
Subject M.14									
28 wks.	69.6	52.4	27.8	15.4	13.1	44.6	11.4	43.8	8.61
32 wks.	70.6	53.3	29.1	15.3	13.2	44.5	11.6	44.5	8.73
40 wks.	72.2	54.4	30.3	17.3	13.2	46.3	12.6	45.5	9.20
48 wks.	74.9	56.8	31.7	16.8	13.6	48.9	12.4	46.4	10.24
52 wks.	76.5	58.5	32.4	—	13.5	—	12.3	46.6	10.51
Subject M.15									
16 wks.	62.5	45.4	24.3	14.8	12.2	41.4	11.7	41.6	5.95
20 wks.	64.5	46.9	25.7	15.3	12.7	42.4	11.9	42.6	7.18
24 wks.	66.3	48.7	26.2	16.6	13.2	44.5	11.7	43.7	7.69
28 wks.	67.5	50.0	26.4	16.1	13.9	46.0	11.7	44.0	8.21
52 wks.	74.9	55.1	30.5	17.9	15.7	49.5	12.5	47.5	10.67
Subject M.16									
28 wks.	64.7	48.0	25.7	15.7	14.4	47.5	12.2	44.9	8.46
32 wks.	66.8	49.5	27.5	16.2	15.4	49.5	12.4	46.0	8.93
36 wks.	67.7	50.2	27.5	16.8	15.4	49.4	12.4	46.6	9.50
40 wks.	69.0	50.5	27.9	16.9	15.5	51.2	12.4	47.2	9.86
44 wks.	69.5	51.2	28.9	17.4	—	—	13.3	47.8	10.55
48 wks.	70.5	53.4	30.0	17.7	16.1	51.8	13.3	48.2	10.96
52 wks.	72.5	53.4	30.1	17.9	16.1	51.1	13.4	48.2	11.16
Subject M.17									
32 wks.	67.9	49.7	26.2	14.8	12.6	44.4	11.1	46.4	8.18
36 wks.	68.7	50.5	26.8	15.6	13.2	—	11.5	46.6	8.52
40 wks.	69.7	50.6	28.0	15.1	13.4	46.5	11.4	47.6	9.17
44 wks.	70.7	52.1	28.2	15.7	13.6	45.6	11.7	48.0	9.12
48 wks.	71.6	52.7	28.5	16.4	13.9	44.9	11.8	48.2	9.25
52 wks.	72.6	53.6	28.8	16.6	14.1	46.4	12.2	48.4	9.32
Subject M.18									
28 wks.	67.1	50.6	26.2	15.1	13.7	42.6	10.9	43.3	7.56
36 wks.	70.1	52.4	27.9	16.1	14.4	44.2	11.2	44.3	8.40
40 wks.	71.2	53.0	28.5	17.0	14.2	46.5	11.9	44.5	8.74
44 wks.	71.4	53.7	29.2	17.5	14.4	46.6	12.2	44.7	9.44
48 wks.	73.3	55.3	29.9	17.8	14.2	47.0	11.9	45.0	9.41
52 wks.	75.3	58.5	31.4	18.0	14.8	48.2	12.1	45.5	9.62

TABLE 3—Continued

Raw data on males

AGE	VERTEX HEIGHT	SUPRA-STERNAL HEIGHT	SYM-PHYSIS HEIGHT	BIA-CROMIAL DIAM-ETER	CHEST BREADTH	CHEST CIRCUM.	BI-CRISTAL DIAM-ETER	HEAD CIRCUM.	WEIGHT
Subject M.19									
28 wks.	65.6	47.9	26.7	14.4	11.7	42.7	10.1	43.2	7.09
32 wks.	67.2	49.7	27.9	—	12.3	42.6	10.8	44.3	8.07
36 wks.	68.0	50.6	28.8	15.3	13.0	43.2	11.1	44.8	8.37
40 wks.	69.4	51.1	28.8	14.5	13.0	44.2	11.5	45.5	8.83
44 wks.	70.6	52.3	29.5	15.9	13.5	45.6	11.1	45.7	8.95
48 wks.	70.9	53.0	30.7	15.4	13.9	44.9	12.2	46.1	9.17
Subject M.20									
32 wks.	70.1	52.3	27.9	17.2	15.0	47.7	12.0	44.5	9.26
36 wks.	71.1	52.7	30.0	17.5	15.0	47.0	12.0	45.0	9.82
44 wks.	73.6	54.6	30.8	17.7	15.5	49.6	12.8	46.9	10.23
48 wks.	75.0	55.7	31.6	18.2	16.3	50.6	13.1	47.4	10.64
52 wks.	76.5	57.8	32.3	18.5	16.7	50.6	13.3	47.7	11.11
Subject M.21									
8 wks.	57.2	41.2	21.5	12.1	12.2	36.9	10.0	38.2	4.64
12 wks.	58.2	42.8	22.4	12.7	—	38.8	10.5	39.8	5.20
16 wks.	61.7	44.3	24.9	13.9	13.7	39.5	10.8	41.0	5.86
56 wks.	75.3	56.0	31.6	17.4	15.5	45.7	12.6	46.2	—
Subject M.22									
8 wks.	61.5	45.5	23.6	12.8	12.0	40.2	9.5	40.2	5.54
12 wks.	64.5	47.2	25.5	14.3	12.6	41.2	9.7	42.0	6.64
16 wks.	67.1	49.2	27.4	14.8	12.7	41.6	10.4	43.0	7.64
56 wks.	82.3	61.8	35.4	18.7	15.9	49.4	13.4	49.4	12.64
Subject M 23.									
8 wks.	60.6	—	23.6	13.1	11.7	35.3	9.7	39.2	4.58
12 wks.	62.2	45.9	24.6	15.0	13.0	38.6	10.6	39.9	5.19
16 wks.	65.2	47.4	26.0	14.2	13.0	39.6	10.6	41.4	6.20
56 wks.	80.6	60.2	33.8	18.4	15.7	49.0	13.2	46.7	10.82
Subject M.24									
4 wks.	54.6	41.0	21.6	11.6	10.8	35.0	8.8	37.8	3.90
6 wks.	56.3	41.8	21.9	13.1	10.7	35.5	9.0	39.0	4.28
8 wks.	58.2	43.2	22.6	13.2	11.3	36.4	9.6	39.8	4.60
12 wks.	60.8	45.0	24.3	13.9	12.2	38.1	9.8	40.9	5.15
16 wks.	61.5	45.4	25.2	14.2	12.4	37.4	10.2	41.0	5.48
56 wks.	76.2	57.1	32.8	17.6	15.1	47.4	12.1	46.4	9.10
Subject M.25									
6 wks.	53.5	39.9	20.8	12.6	10.7	34.7	9.4	36.9	4.10
8 wks.	55.2	41.4	21.6	13.0	11.4	35.6	9.8	37.5	4.35
12 wks.	58.2	42.7	22.6	12.4	11.4	37.4	10.3	38.9	5.10
16 wks.	61.3	45.3	23.9	13.3	12.2	39.3	10.8	40.1	5.91
56 wks.	77.1	58.6	32.4	17.2	16.4	48.8	13.3	46.4	11.09

TABLE 3—Continued

Raw data on males

AGE	VERTEX HEIGHT	SUPRA-STERNAL HEIGHT	SYM-PHYSIS HEIGHT	BIA-CROMIAL DIAM-ETER	CHEST BREADTH	CHEST CIRCUM.	BI-CRISTAL DIAM-ETER	HEAD CIRCUM.	WEIGHT
Subject M.26									
8 wks.	60.3	45.2	25.6	14.6	12.0	40.4	10.0	40.6	5.70
12 wks.	64.4	47.7	26.8	14.0	12.6	41.0	10.7	42.0	6.79
16 wks.	67.4	50.9	—	15.4	13.2	43.6	10.7	43.4	7.42
56 wks.	83.6	63.8	37.4	19.7	16.9	48.7	14.0	48.7	12.86
Subject M.27									
36 wks.	70.3	51.9	29.2	14.7	12.9	42.5	11.5	46.3	8.10
40 wks.	72.2	53.1	30.2	16.0	13.4	45.3	11.8	47.0	8.52
44 wks.	73.8	54.0	30.2	16.1	13.6	45.0	12.1	47.4	8.88
52 wks.	76.5	57.4	31.6	16.4	14.7	45.4	12.3	47.7	9.14
Subject M.28									
32 wks.	70.7	52.1	28.4	15.6	13.5	44.0	11.2	44.0	7.68
36 wks.	71.4	52.8	29.2	15.7	13.7	44.5	11.5	44.8	7.92
40 wks.	72.7	54.0	29.3	17.0	14.4	49.1	11.6	45.2	8.75
44 wks.	74.8	55.6	30.3	17.2	14.7	49.2	11.7	45.6	9.12
Subject M.29									
8 wks.	58.6	43.8	21.6	13.3	11.7	40.4	9.6	38.8	5.60
12 wks.	61.4	46.0	23.4	14.2	12.5	40.6	10.6	40.1	6.25
16 wks.	63.6	47.2	24.3	13.6	12.7	43.2	10.9	40.9	6.98
Subject M.30									
8 wks.	57.4	41.2	21.7	11.8	10.6	36.2	8.6	40.1	—
12 wks.	61.2	45.2	23.2	11.6	11.2	36.6	9.8	40.7	—
16 wks.	63.7	46.4	24.7	13.4	12.2	38.9	10.4	(38.7)	—
Subject M.31									
8 wks.	55.2	40.4	21.3	12.6	11.4	37.4	9.6	(36.2)	5.04
12 wks.	58.2	42.5	21.5	13.6	13.2	40.8	10.7	40.2	5.90
16 wks.	62.2	44.8	24.8	13.4	13.2	—	11.6	41.8	6.81
Subject M.32									
24 wks.	67.0	49.4	25.8	15.8	12.3	43.4	11.4	42.0	7.88
28 wks.	68.3	50.1	26.5	16.8	12.8	45.1	11.7	(44.5)	8.72
32 wks.	69.1	50.7	27.9	16.7	14.0	45.1	12.9	43.7	9.34

TABLE 4

Raw data on females

AGE	VERTEX HEIGHT	SUPRA-STERNAL HEIGHT	SYM-PHYSIS HEIGHT	BIA-CROMIAL DIAM-ETER	CHEST BREADTH	CHEST CIRCUM.	BI-CRISTAL DIAM-ETER	HEAD CIRCUM.	WEIGHT
Subject F.1									
12 wks.	61.2	45.6	24.6	13.8	12.4	—	10.4	—	6.10
16 wks.	63.7	47.7	25.9	15.0	12.9	42.5	10.9	40.6	6.89
20 wks.	66.1	49.4	28.0	15.8	13.1	44.9	11.1	41.6	7.53
24 wks.	69.2	51.7	28.6	16.2	13.6	45.0	11.5	42.3	7.87
28 wks.	70.0	52.2	29.7	16.5	13.6	44.0	11.6	43.0	7.96
32 wks.	71.8	52.9	30.4	17.3	14.4	44.4	11.6	43.3	8.40
36 wks.	72.2	53.4	30.9	17.6	14.6	45.8	11.9	43.6	8.90
40 wks.	75.0	57.0	31.1	17.6	14.5	46.8	11.8	44.2	9.54
44 wks.	76.3	57.1	32.4	17.6	14.5	47.2	12.2	44.5	9.72
48 wks.	77.6	57.6	33.5	17.9	15.6	47.8	12.6	44.8	9.83
52 wks.	78.6	58.9	34.5	18.4	15.8	48.6	12.6	45.1	10.38
60 wks.	81.4	61.7	35.5	18.6	15.8	50.5	13.1	46.0	10.95
80 wks.	85.2	64.5	38.4	18.2	16.1	(48.8)	13.4	46.1	10.24
2 yrs.	90.1	68.8	40.8	20.3	16.7	51.4	14.2	46.9	12.12
3 yrs.	97.8	74.9	46.0	19.4	16.8	51.9	15.6	47.4	—
4 yrs.	105.2	82.3	50.0	22.4	17.4	53.8	16.5	48.5	16.72
5 yrs.	111.8	87.8	54.4	23.9	17.9	53.7	17.3	48.5	18.14
6 yrs.	117.6	93.0	59.9	24.6	18.8	54.5	18.1	49.0	19.94
7 yrs.	124.3	99.1	62.9	26.3	19.3	—	18.9	—	22.31
8 yrs.	128.3	102.5	65.1	27.0	19.7	58.2	19.4	49.5	23.41
9 yrs.	132.1	105.9	67.9	27.4	20.2	59.0	20.4	50.0	25.11
10 yrs.	136.1	108.9	—	28.7	20.7	—	—	50.2	26.36
Subject F.2									
8 wks.	59.7	43.5	23.5	12.7	12.0	35.8	9.4	39.5	5.16
12 wks.	61.7	—	—	13.4	12.6	38.4	9.8	40.6	5.65
16 wks.	64.4	47.4	25.5	14.0	13.4	39.8	10.4	41.3	6.49
20 wks.	67.6	50.3	26.8	14.8	—	42.1	10.6	42.7	6.81
24 wks.	68.3	50.9	27.2	16.2	13.8	43.0	11.1	43.2	7.54
28 wks.	70.0	51.8	29.0	16.0	14.2	43.5	11.5	44.4	8.08
32 wks.	71.2	52.5	29.5	16.2	14.5	44.7	11.6	45.7	8.85
36 wks.	72.1	54.8	29.8	16.8	14.7	47.8	11.9	46.4	9.24
40 wks.	73.5	55.0	30.6	16.8	15.3	48.1	12.2	46.6	9.63
44 wks.	74.6	55.8	31.7	16.9	15.6	49.9	12.3	47.2	9.72
48 wks.	76.2	56.5	32.0	17.8	16.4	49.3	12.6	47.3	10.64
52 wks.	76.5	57.6	32.8	18.6	16.3	(47.6)	12.6	47.9	10.19
56 wks.	78.2	59.2	34.2	18.7	15.9	49.6	12.8	48.4	10.97
60 wks.	79.2	59.6	34.3	18.8	16.6	51.4	12.8	48.3	11.06
2 yrs.	88.4	67.4	39.9	20.6	18.3	51.1	14.2	50.4	12.46
3 yrs.	95.8	73.9	45.4	21.3	—	53.8	15.3	50.8	13.98
4 yrs.	104.9	81.0	51.1	23.0	18.6	56.0	17.0	51.4	17.54
5 yrs.	110.0	87.5	54.9	24.1	19.4	57.4	17.8	52.0	19.30
6 yrs.	119.2	94.2	60.2	26.3	19.5	58.8	18.5	51.4	20.82
7 yrs.	124.2	98.5	63.3	27.7	20.3	60.8	(17.3)	51.5	23.96
8 yrs.	131.5	105.4	67.6	28.7	21.5	63.0	20.2	52.2	28.07
9 yrs.	136.6	110.1	71.1	29.7	22.4	65.8	20.7	53.8	31.36
10 yrs.	142.5	115.8	75.3	30.4	23.3	73.9	22.4	55.2	42.50

TABLE 4—Continued

Raw data on females

AGE	VERTEX HEIGHT	SUPRA-STERNAL HEIGHT	SYM-PHYSIS HEIGHT	BIA-CROMIAL DIAM-ETER	CHEST BREADTH	CHEST CIRCUM.	BI-CRISTAL DIAM-ETER	HEAD CIRCUM.	WEIGHT
Subject F.3									
12 wks.	60.0	—	24.3	12.3	12.2	36.4	10.2	38.9	5.02
16 wks.	61.7	46.2	25.7	14.4	11.3	39.0	9.7	39.5	5.22
20 wks.	63.4	47.9	26.6	14.2	12.6	41.4	10.4	40.6	6.06
24 wks.	65.5	48.5	28.1	14.6	12.7	41.2	10.4	41.4	6.29
28 wks.	67.0	50.0	28.0	15.1	12.6	41.1	10.5	42.0	6.61
32 wks.	67.7	51.0	28.8	15.7	13.2	43.6	10.7	42.6	6.96
36 wks.	69.2	51.5	29.8	15.7	13.4	42.8	10.8	43.2	7.03
40 wks.	71.1	53.2	31.2	16.0	13.6	42.6	12.0	43.5	7.37
44 wks.	71.4	53.5	31.3	16.3	13.9	42.5	12.0	44.2	7.57
48 wks.	73.4	56.0	32.0	16.9	14.0	44.1	11.5	44.0	7.77
52 wks.	75.0	56.6	32.8	17.0	14.3	45.0	11.8	44.4	8.36
56 wks.	76.2	58.1	33.9	17.2	14.6	45.3	11.8	44.8	8.59
60 wks.	77.5	58.7	34.6	17.7	14.6	46.0	12.2	45.5	8.81
80 wks.	82.0	62.3	37.0	18.2	14.9	45.8	12.7	46.0	9.00
2 yrs.	85.3	65.7	39.3	18.7	15.9	46.8	13.4	47.0	9.67
3 yrs.	95.1	73.8	45.4	20.3	16.3	49.0	15.1	47.6	12.19
4 yrs.	103.2	80.9	50.5	21.7	16.6	51.4	16.0	48.7	14.21
5 yrs.	111.0	87.7	56.2	22.8	17.1	52.0	16.6	48.8	15.64
6 yrs.	118.6	94.5	60.6	24.5	17.8	53.6	17.3	49.3	17.62
7 yrs.	123.9	99.1	64.2	25.4	18.5	54.5	18.3	49.9	19.90
8 yrs.	129.8	104.5	68.5	25.7	19.2	57.4	19.0	50.9	22.61
9 yrs.	134.5	108.9	70.9	27.8	19.8	58.5	19.6	51.0	25.00
10 yrs.	139.7	113.8	74.9	26.5	20.4	62.9	20.6	50.9	28.41
Subject F.4									
6 wks.	52.2	38.7	20.6	12.4	11.2	34.2	9.2	36.4	3.80
8 wks.	54.7	41.0	22.1	12.2	11.8	—	8.6	36.9	4.00
12 wks.	56.6	42.1	22.8	12.4	12.2	38.6	9.2	37.9	4.73
16 wks.	58.4	43.5	24.1	13.4	12.4	40.3	10.2	40.2	5.26
20 wks.	60.5	45.0	24.2	13.8	13.2	41.6	10.6	41.5	5.94
24 wks.	62.3	46.3	25.1	14.2	13.4	42.2	10.7	42.0	—
28 wks.	62.1	46.5	26.2	14.3	13.4	43.9	11.0	43.7	6.87
32 wks.	64.3	47.5	26.8	14.5	13.5	44.6	11.0	44.0	7.54
36 wks.	66.8	49.9	27.6	14.9	14.2	46.4	11.2	45.1	7.92
40 wks.	67.7	50.7	27.9	15.6	14.3	47.3	11.6	45.8	8.42
44 wks.	69.6	52.4	29.1	16.7	14.6	47.4	11.7	45.8	8.43
48 wks.	68.9	51.7	28.7	16.6	15.0	48.9	11.8	46.4	9.14
52 wks.	—	—	—	—	—	49.9	—	—	—
56 wks.	71.7	54.0	30.9	17.5	15.0	52.4	12.3	47.0	10.14
60 wks.	73.5	55.3	32.1	18.4	16.1	51.0	12.5	47.6	10.45
80 wks.	78.8	60.1	—	19.2	16.8	51.4	13.3	48.2	11.33
2 yrs.	83.2	64.5	38.8	20.0	17.3	52.7	14.2	48.5	12.93
3 yrs.	91.8	71.2	43.6	21.1	17.4	54.2	15.5	49.3	14.49
4 yrs.	99.1	78.4	48.3	22.8	18.1	55.8	16.7	50.0	16.26
5 yrs.	106.2	84.5	52.7	24.4	19.0	57.7	17.9	50.3	18.92
6 yrs.	112.1	89.5	56.6	25.7	19.7	58.7	18.7	51.1	21.01
7 yrs.	117.4	94.4	59.4	27.1	20.2	63.0	19.4	51.0	25.26
8 yrs.	123.1	99.9	63.6	28.2	21.3	65.9	20.5	51.5	29.10
9 yrs.	127.5	103.2	66.5	28.4	22.2	66.5	20.6	52.0	30.87
10 yrs.	132.6	108.0	69.9	29.6	22.4	71.2	22.1	52.7	34.19

TABLE 4—Continued

Raw data on females

AGE	VERTEX HEIGHT	SUPRA-STERNAL HEIGHT	SYM-PHYSIS HEIGHT	BIA-CROMIAL DIAM-ETER	CHEST BREADTH	CHEST CIRCUM.	BI-CRISTAL DIAM-ETER	HEAD CIRCUM.	WEIGHT
Subject F.5									
16 wks.	60.8	44.6	23.0	12.3	11.0	39.4	10.1	39.5	6.15
20 wks.	62.6	47.0	24.1	13.2	10.9	42.0	10.3	40.5	7.16
28 wks.	65.8	48.4	25.7	13.8	11.7	44.3	10.9	42.4	8.22
32 wks.	68.7	50.8	26.7	14.8	12.5	44.1	11.2	43.0	8.46
44 wks.	71.2	53.3	30.4	15.9	12.7	46.1	12.1	44.1	9.20
48 wks.	72.2	54.3	30.7	15.2	12.6	46.9	12.1	44.6	9.54
52 wks.	73.2	54.3	30.3	15.5	—	47.4	11.4	—	9.66
56 wks.	74.1	55.1	32.3	16.1	14.2	47.8	12.1	44.6	9.96
Subject F.6									
16 wks.	60.1	44.4	24.6	12.6	11.5	39.9	10.1	39.5	5.74
20 wks.	62.4	46.2	25.4	13.7	12.0	40.5	—	40.5	6.18
28 wks.	65.4	48.5	27.1	14.4	12.6	42.0	10.4	41.8	6.62
40 wks.	69.9	52.1	26.6	15.6	14.2	44.2	11.5	43.4	8.12
48 wks.	72.3	53.4	29.8	15.9	14.1	46.3	11.8	44.3	8.56
52 wks.	73.0	54.6	29.9	16.1	14.6	46.5	11.2	44.4	8.79
Subject F.7									
20 wks.	62.6	45.9	24.9	13.2	11.6	39.9	10.1	41.1	6.29
24 wks.	63.7	47.6	25.7	13.6	12.2	40.8	10.4	41.5	6.55
36 wks.	67.4	49.0	27.5	15.1	11.7	44.5	10.9	42.6	7.57
44 wks.	69.7	51.6	28.0	14.8	12.9	44.6	12.4	43.5	8.60
48 wks.	70.8	52.5	29.2	16.3	12.7	45.5	11.6	44.0	9.01
52 wks.	72.0	53.0	29.6	15.4	13.2	45.2	11.6	44.0	8.85
56 wks.	72.7	53.6	30.7	16.2	13.1	45.0	12.4	44.5	9.20
Subject F.8									
16 wks.	61.6	45.7	24.4	13.9	11.8	43.1	10.5	41.4	6.19
20 wks.	64.5	47.9	26.0	14.8	12.2	42.1	10.6	42.1	6.71
36 wks.	69.4	51.4	29.3	14.6	(14.2)	44.2	11.8	44.4	8.52
40 wks.	70.4	52.0	30.1	14.8	13.2	46.4	12.1	44.4	8.68
52 wks.	73.8	55.1	31.1	15.6	13.6	45.1	12.4	45.3	8.99
Subject F.9									
16 wks.	60.4	45.4	24.0	13.8	12.6	40.4	10.5	38.9	5.75
20 wks.	62.4	46.8	25.8	13.7	12.4	42.4	10.6	39.9	6.45
24 wks.	64.7	48.6	26.5	14.1	12.1	43.4	10.6	40.7	7.04
28 wks.	65.9	49.4	27.1	14.3	13.1	44.1	11.0	41.0	7.55
48 wks.	70.8	51.5	29.9	16.1	14.1	46.6	12.1	43.5	9.12
52 wks.	71.6	54.6	30.7	16.7	14.6	47.9	12.2	44.3	9.67
Subject F.10									
20 wks.	62.9	46.1	25.6	13.5	11.6	39.5	9.2	42.2	6.32
24 wks.	64.9	48.4	26.2	—	11.4	41.0	10.4	43.0	6.95
28 wks.	65.8	48.2	28.5	14.3	12.4	41.6	10.7	43.6	7.28
44 wks.	70.7	51.9	29.3	14.8	12.4	41.5	10.3	45.2	8.10
48 wks.	71.8	—	30.0	15.8	13.4	42.3	11.2	46.1	8.08
52 wks.	72.4	54.3	30.5	15.8	12.7	43.5	11.6	46.0	8.49

TABLE 4—Continued

Raw data on females

AGE	VERTEX HEIGHT	SUPRA-STERNAL HEIGHT	SYM-PHYSIS HEIGHT	BIA-CROMIAL DIAM-ETER	CHEST BREADTH	CHEST CIRCUM.	BI-CRISTAL DIAM-ETER	HEAD CIRCUM.	WEIGHT
Subject F.11									
24 wks.	66.0	48.9	26.8	14.4	12.8	40.1	10.9	41.6	6.64
28 wks.	67.0	49.9	27.9	15.0	13.4	41.9	11.3	41.9	6.73
32 wks.	69.3	51.4	27.9	15.2	13.7	43.2	11.5	42.3	7.40
36 wks.	69.4	51.5	29.3	16.7	14.1	43.8	11.8	42.7	7.48
40 wks.	71.2	52.7	29.7	16.9	14.1	43.9	12.2	43.0	7.77
44 wks.	71.8	52.8	30.1	17.4	14.2	43.9	12.3	43.4	7.82
48 wks.	73.1	54.2	30.5	17.5	14.4	45.5	12.6	43.7	8.36
52 wks.	74.1	55.0	31.3	17.5	14.2	44.4	12.6	43.7	8.29
Subject F.12									
24 wks.	62.3	45.6	25.6	14.8	13.7	43.2	11.5	41.9	7.49
28 wks.	65.2	47.7	25.6	15.2	13.6	45.9	11.8	42.5	8.00
32 wks.	66.2	48.5	25.9	15.3	13.6	44.5	11.9	43.2	8.46
36 wks.	67.2	49.7	26.7	15.7	14.1	47.9	12.1	44.1	8.86
40 wks.	68.7	50.6	27.9	16.7	14.6	48.0	12.2	44.9	9.20
44 wks.	70.0	51.6	28.6	17.1	14.7	48.5	12.4	45.3	9.49
52 wks.	72.5	54.7	30.2	17.1	14.8	49.0	12.4	46.0	10.00
56 wks.	73.8	56.3	30.8	18.1	15.1	49.1	12.5	46.3	10.13
80 wks.	80.6	61.1	34.5	17.9	15.9	50.1	13.5	47.4	11.13
2 yrs.	85.6	65.4	38.2	20.3	16.5	50.8	14.7	48.4	12.31
Subject F.13									
32 wks.	69.6	52.8	29.2	15.7	14.4	42.3	11.6	44.2	7.59
36 wks.	71.1	52.7	29.9	16.6	—	44.4	12.2	44.9	7.88
40 wks.	72.5	54.2	30.6	17.2	14.6	44.8	12.5	45.3	8.21
44 wks.	74.2	55.7	30.8	17.7	15.3	45.1	12.5	45.6	8.67
48 wks.	75.0	56.2	32.1	18.2	15.6	45.8	12.5	46.0	8.83
52 wks.	76.6	57.5	32.8	18.4	15.8	46.6	12.8	46.2	9.30
56 wks.	77.9	58.5	33.3	18.2	15.2	46.5	12.8	46.5	9.25
Subject F.14									
28 wks.	66.0	49.7	27.6	14.6	12.2	41.5	9.8	41.5	6.34
36 wks.	69.0	52.2	29.1	15.0	12.7	42.4	10.6	42.4	5.80
40 wks.	70.8	53.1	30.1	—	12.7	43.4	10.6	42.9	7.18
44 wks.	70.9	53.8	30.7	15.0	13.2	42.2	10.6	42.2	6.94
48 wks.	72.1	54.8	31.1	15.6	13.2	42.9	11.2	43.4	7.94
52 wks.	73.0	54.9	32.0	16.0	13.2	44.2	11.7	44.2	7.84
Subject F.15									
32 wks.	67.9	50.4	27.0	15.8	14.2	43.5	11.1	43.0	7.59
36 wks.	69.7	51.6	27.6	16.8	14.2	—	11.7	43.4	8.00
40 wks.	70.7	53.0	28.5	17.2	15.0	44.2	12.1	44.2	8.74
44 wks.	72.6	54.6	30.1	17.6	14.6	46.4	12.1	44.7	9.04
48 wks.	73.7	55.4	30.6	18.3	14.6	47.2	12.3	44.9	9.32
52 wks.	75.8	57.3	31.6	18.8	15.4	47.5	12.4	45.5	9.64
56 wks.	77.2	57.7	33.1	—	—	—	—	—	10.14

TABLE 4—Continued

Raw data on females

AGE	VERTEX HEIGHT	SUPRA-STERNAL HEIGHT	SYM-PHYSIS HEIGHT	BIA-CROMIAL DIAM-ETER	CHEST BREADTH	CHEST CIRCUM.	BI-CRISTAL DIAM-ETER	HEAD CIRCUM.	WEIGHT
Subject F.16									
20 wks.	59.9	44.2	23.8	13.9	10.8	40.9	8.8	39.9	5.19
24 wks.	62.4	45.9	24.9	14.3	11.9	41.7	10.5	40.8	6.20
28 wks.	64.4	47.6	25.8	15.0	12.9	43.6	11.4	42.1	7.03
52 wks.	72.5	54.5	31.4	16.5	13.4	46.8	12.6	44.4	8.90
Subject F.17									
20 wks.	60.7	45.2	24.2	15.4	12.6	—	12.0	40.0	6.26
24 wks.	63.2	47.3	25.0	15.5	13.4	42.8	11.1	40.9	6.70
28 wks.	63.8	46.4	26.2	(14.4)	13.3	44.0	10.8	41.7	7.15
52 wks.	71.0	53.3	29.6	16.5	14.2	45.9	12.2	44.9	9.14
Subject F.18									
24 wks.	66.5	49.3	28.7	14.2	11.7	41.2	11.1	42.3	7.08
28 wks.	66.0	48.4	—	14.3	11.9	42.7	10.7	43.1	7.46
32 wks.	68.5	50.7	31.5	—	12.5	44.5	12.5	44.0	8.06
52 wks.	73.5	54.5	—	—	—	—	—	45.6	8.75
Subject F.19									
8 wks.	58.5	42.8	23.4	13.9	11.8	36.8	10.0	37.6	4.48
12 wks.	61.8	46.0	25.1	13.7	12.2	39.2	10.6	39.0	5.35
16 wks.	64.4	47.6	25.9	13.2	13.9	40.8	11.2	38.9	6.20
56 wks.	80.2	59.5	34.6	19.4	15.8	48.4	13.2	45.4	10.75
Subject F.20									
8 wks.	57.4	42.4	23.2	12.5	10.8	36.5	10.0	37.8	4.92
12 wks.	59.4	44.2	24.4	13.5	11.7	37.4	9.8	39.3	5.37
16 wks.	59.8	44.7	25.5	13.0	11.6	38.6	10.2	40.5	5.54
56 wks.	75.3	56.0	31.4	17.0	14.7	47.0	12.0	47.2	10.26
Subject F.21									
8 wks.	55.4	41.4	20.0	11.7	11.0	38.5	10.4	38.8	4.70
12 wks.	58.0	42.4	22.0	12.8	12.3	38.6	10.2	40.2	5.62
16 wks.	61.4	45.1	22.5	14.6	12.5	40.0	10.2	41.5	6.17
56 wks.	78.7	59.1	32.2	17.7	15.2	48.6	12.9	46.9	10.29
Subject F.22									
8 wks.	54.1	40.0	20.9	12.4	9.6	32.4	8.9	36.5	3.70
12 wks.	55.2	41.8	22.4	11.6	10.8	35.0	9.2	37.5	4.25
16 wks.	59.4	44.0	23.0	13.7	11.9	38.4	10.1	38.9	5.37
56 wks.	72.5	54.6	30.3	17.8	14.4	44.2	12.7	43.9	8.72
Subject F.23									
20 wks.	63.7	47.5	25.7	15.7	11.3	40.9	10,8	40.8	6.00
24 wks.	65.2	48.5	26.1	15.3	11.9	42.9	11.7	42.1	6.76
28 wks.	68.0	51.2	27.8	16.3	11.4	44.4	11.1	42.9	7.68
52 wks.	76.5	57.7	33.5	17.7	—	48.4	13.8	45.9	10.25

TABLE 4—Continued

Raw data on females

AGE	VERTEX HEIGHT	SUPRA-STERNAL HEIGHT	SYM-PHYSIS HEIGHT	BIA-CROMIAL DIAM-ETER	CHEST BREADTH	CHEST CIRCUM.	BI-CRISTAL DIAM-ETER	HEAD CIRCUM.	WEIGHT
Subject F.24									
8 wks.	54.9	41.0	21.5	12.4	10.6	36.3	9.5	36.7	4.45
12 wks.	57.7	43.0	22.6	12.8	10.9	37.3	10.2	37.2	5.20
16 wks.	60.3	44.5	24.2	14.2	12.3	39.7	10.6	38.3	5.42
56 wks.	76.6	58.7	32.9	17.7	14.8	46.2	12.7	43.5	9.51
Subject F.25									
32 wks.	64.3	47.3	25.8	14.6	12.7	42.6	10.6	41.9	6.88
36 wks.	66.1	48.7	26.3	15.0	13.3	42.8	10.8	42.6	7.55
40 wks.	67.7	50.2	26.6	15.3	13.4	43.9	10.9	42.8	7.72
44 wks.	68.4	50.5	27.5	15.6	13.5	44.6	11.4	43.5	8.34
48 wks.	70.2	52.5	28.4	16.7	13.6	45.4	11.6	43.8	8.94
52 wks.	72.1	54.0	29.1	17.1	13.8	46.8	11.8	44.4	9.33
Subject F.26									
32 wks.	69.0	52.3	26.1	16.3	12.3	—	11.2	41.6	7.78
36 wks.	69.9	52.3	29.5	16.3	13.1	42.4	11.2	42.0	7.96
44 wks.	71.8	53.4	29.9	16.6	13.4	43.4	11.7	42.9	8.43
48 wks.	72.9	55.5	30.3	17.2	13.4	44.0	12.3	43.0	8.74
52 wks.	74.0	56.2	30.5	17.6	13.4	45.0	12.4	43.3	8.72
Subject F.27									
36 wks.	68.9	51.9	30.1	15.9	13.8	42.8	11.5	42.4	7.17
40 wks.	70.1	52.3	29.9	16.3	—	41.6	11.6	42.4	7.27
44 wks.	71.0	52.5	30.3	16.6	14.1	43.5	11.7	42.9	7.61
48 wks.	72.9	54.3	31.2	16.6	14.1	44.6	12.3	42.9	7.92
52 wks.	73.6	54.8	31.6	17.6	14.6	46.3	12.6	43.7	8.10
Subject F.28									
20 wks.	64.9	48.1	26.3	16.3	12.2	45.4	11.0	39.9	7.30
24 wks.	66.7	50.1	27.1	15.7	12.9	(43.9)	10.5	40.8	7.87
28 wks.	68.5	51.5	27.9	16.3	13.4	46.9	10.6	41.0	8.39
36 wks.	70.9	53.9	29.6	16.2	12.9	46.1	11.3	41.7	8.36
40 wks.	72.2	54.6	29.8	16.3	13.6	47.6	11.4	42.3	9.02
52 wks.	75.0	—	—	—	—	—	—	—	9.94
Subject F.29									
32 wks.	70.6	52.9	28.9	16.1	13.6	44.9	11.1	47.6	9.17
36 wks.	72.1	53.3	30.1	—	13.6	44.6	11.1	48.5	9.32
40 wks.	73.1	54.2	—	16.8	14.0	45.0	11.4	48.8	9.30
48 wks.	75.4	56.3	31.6	16.8	14.7	46.4	11.5	49.8	10.10
Subject F.30									
24 wks.	66.1	48.9	27.1	15.9	12.9	44.6	10.9	42.4	8.06
28 wks.	67.8	50.9	28.5	16.1	13.7	45.9	11.2	43.3	8.33
32 wks.	69.0	51.6	29.1	15.8	(12.1)	45.5	11.8	43.6	8.37

TABLE 4—Continued

Raw data on females

AGE	VERTEX HEIGHT	SUPRA-STERNAL HEIGHT	SYM-PHYSIS HEIGHT	BIA-CROMIAL DIAM-ETER	CHEST BREADTH	CHEST CIRCUM.	BI-CRISTAL DIAM-ETER	HEAD CIRCUM.	WEIGHT
Subject F.31									
8 wks.	60.7	44.3	23.7	14.0	11.0	37.3	10.3	38.3	4.90
12 wks.	62.8	46.8	25.1	14.0	11.9	38.2	11.0	39.4	5.50
16 wks.	65.2	48.0	26.1	14.8	13.6	41.6	11.5	40.1	6.64
Subject F.32									
24 wks.	67.7	49.8	28.4	15.7	12.4	46.1	11.6	41.3	7.20
28 wks.	68.4	51.3	28.5	15.3	(13.8)	46.2	11.4	42.0	7.58
32 wks.	70.2	52.7	30.2	15.8	12.5	(44.1)	12.3	42.6	8.08
Subject F.33									
8 wks.	58.4	42.2	23.3	13.4	11.6	39.2	9.4	39.4	5.10
12 wks.	61.2	46.0	24.6	12.9	10.4	38.6	10.0	40.4	6.08
16 wks.	64.2	47.2	25.7	13.6	13.2	42.3	10.6	41.9	7.18
Subject F.34									
8 wks.	54.2	39.2	21.6	10.6	10.5	35.2	8.9	38.6	4.20
12 wks.	55.2	39.6	22.4	12.3	10.4	35.9	9.3	39.9	4.64
16 wks.	57.5	41.7	23.3	12.2	11.6	37.6	9.7	41.2	5.28
Subject F.35									
8 wks.	54.0	39.9	20.8	12.7	11.1	36.4	8.9	38.4	4.30
12 wks.	57.2	42.9	22.0	13.0	12.5	39.0	10.0	38.7	5.20
16 wks.	59.6	44.3	23.4	12.6	12.3	40.4	10.6	40.2	6.20
Subject F.36									
40 wks.	69.7	52.5	28.7	15.4	12.8	44.4	11.6	43.1	8.54
44 wks.	70.3	53.0	29.1	—	13.8	45.5	11.8	43.5	9.10
48 wks.	72.0	54.1	29.4	17.0	14.3	46.8	12.4	44.1	9.88

MEASUREMENTS

All measurements except those at 8 and 12 weeks were made by the author. Those at 8 and 12 weeks were partly made by the author and partly by one other person.

1. *Vertex height* (*soles — vertex*). The infant is stretched horizontally on a measuring board with the soles of the feet held in contact with one end of the board; the anthropometer arm is brought down from the other end till it touches the vertex of the head.

2. *Suprasternal height* (*soles — suprasternal*). The same, with the anthropometer arm brought to the suprasternal notch.

3. *Symphysis height* (*soles — pubes*). The same, with anthropometer arm brought to crease above pubes, or to upper border of symphysis pubis.

4. *Biacromial diameter.* Taken while the infant still lies on the board. The infant must be relaxed and with the shoulders drawn neither forward nor back. Spreading calipers.

5. *Chest breadth.* Taken similarly at the nipple level; the point intermediate between minimum and maximum is taken if the child is crying. Spreading calipers.

6. *Chest circumference.* Taken at the nipple level.

7. *Bicristal diameter.* Taken between the most lateral portions of the iliac cress. Spreading calipers.

8. *Head circumference.* Taken round the most prominent portion of the occiput, and the glabella.

REFERENCES

1. GESELL, A. and H. THOMPSON. The Psychology of Early Growth. Macmillan (*New York*), 1938.

2. THOMPSON, H. Body proportions in the growing infant. *Growth*, 2: 1, 1938.

3. TANNER, J. M. Some notes on the reporting of growth data. *Human Biology*, 23: 93, 1951.

4. DAVENPORT, C. B. Bodily growth of babies during the first postnatal year. *Contrib. to Embryol.*, 27: 271, 1938.

SOME NOTES ON THE REPORTING OF GROWTH DATA

BY J. M. TANNER

Sherrington School of Physiology,

St. Thomas's Hospital Medical School, London

INTRODUCTION

AMONGST students of human growth and development one frequently hears discussed the relative advantages of longitudinal and cross-sectional growth studies, and the general opinion that both have their merits usually secures the confident agreement of all concerned. It is less often, however, that the precise nature of these merits is set forth, with the necessary inferences as to the statistical handling of each type of data. Indeed, the need for a comprehensive account of the biometrical handling of growth data is glaringly obvious. Its satisfaction would certainly be a formidable undertaking, to which the present paper makes no pretense. What is given here is little more than a series of notes arising from my own difficulties when faced with the extraction of information from a mass of measurements. But at least they deal with the problem and may stimulate some biometrically-minded biologist with a mathematical statistician at his elbow to attempt a real synthesis of views and methods.

In one section of the field these notes do approach a complete account, at least of the simpler aspects; the appearance of Patterson's paper (34) on sampling on successive occasions with partial replacement has enabled me to rewrite extensively some sections of Part 1, dealing with the reporting of a single measurement, such as stature. I am most grateful to him for allowing me to read his manuscript prior to publication, and to Mr. M. J. R. Healy, who put me in touch with him, and contributed himself a great many helpful criticisms.

These notes are divided into two parts. Part 1, dealing with the single measurement, re-emphasises the importance of reporting longi-

tudinal data [1] in a different way from cross-sectional, and discusses the technique needed to get the maximum information out of the data from mixed longitudinal series, and from a long series of measurements on a single individual. It is still commonplace to see mixed longitudinal series reported in a purely cross-sectional manner, the mean values at age 7, 8, etc., being given; and subsequently in the same article to encounter a statement about "rate of growth" estimated as the mean at 8 minus the mean at 7.[2] Shuttleworth (1) examined this fallacy at some length and with the great clarity in 1937, and Wilson (2) later dealt with it in more mathematical terms. Yet it seems that Dr. Thompson's paper (3), immediately preceding this one, is the first to report actual increments and their variability in a mixed longitudinal series for measurements other than height and weight.

Part 2 deals with the simultaneous reporting of two or more measurements. The growth of one part of the body relative to another has been much written about, and often characterised amongst anthropologists by a ratio index, such as trunk length/leg length. Zoologists, on the other hand, have more often used the allometry equation introduced by Huxley and others, wherein the multiplicative velocities of the two

[1] I have used a terminology as follows. There are three types of series: *cross-sectional*, where each child is measured once only, and all the children at age 8 differ from all those at age 7; *pure longitudinal*, where every child is measured at every age, and where all the children at age 8 are the same as the children at age 7; *mixed longitudinal*, where some or all of the children are measured at least twice, but where some or all fail to last the entire course from entry age to maturity—here some of the children at age 8 are the same as at age 7, but some are new entrants. Pure longitudinal series can as a rule only be arrived at by taking from a mixed longitudinal series those cards that have been successfully completed, and often their number is distressingly small. For this reason, mixed longitudinal series have a great importance, and yet it is in the reporting of these that confusion seems still to occur.

[2] The force of Shuttleworth's (1) remarks on this topic has not diminished during the last decade. "The importance of repeated measurements on the same children," he wrote, "has been urged for so long and by so many as to bcome the first article of faith among students of child development. Nevertheless, the multitude of longitudinal studies which have yielded only cross-sectional findings suggest that this faith has been degenerating to the level of pure dogma. . . . In longitudinal studies the emphasis of the analysis should fall on the growth increments rather than on gross dimensions. . . . Thousands of dollars have been spent on the collection of longitudinal data, hundreds of dollars have been spent on cross-sectional problems, and only dollars have been devoted to the developmental aspects of the data."

dimensions are compared. Both these methods are discussed and the relations between them outlined. There is a brief allusion to the possibilities of dealing with more than two measurements simultaneously, and finally a few remarks on the relationship between allometry during growth and the factor analysis of adult physique.

Throughout I have endeavoured to present the argument in such a way that it is intelligible to the biologist, pediatrician or anthropologist who is not a trained biometrician. I have assumed a knowledge of only very simple biometric procedures such as are implied by the words *mean, standard deviation, variance, standard error, regression* and *correlation coefficient.* Some of the algebraic expressions may be lengthy, but this does not mean they are complicated, and in an appendix I have given fully worked out numerical examples of all of them. A short study of these workings in the appendix should enable a technician computer to work out all the various statistics detailed below.

PART ONE: THE SINGLE MEASUREMENT

The hypotheses

The way in which data of a single measurement, such as stature, should be reported depends on the use to which they are going to be put. Data are collected in order to test hypotheses, and three possible hypotheses are as follows:

A. That child P is abnormal in stature, i. e. that the chance of P belonging to the group G of children of which the standardising data is a sample, is very small;

B. That child P is growing abnormally fast, i. e. that the chance of P's increment between t_1 and t_2 years belonging to the population of increments of the group G is very small;

C. That child P is growing with an abnormal acceleration, or (a more probable clinical event) that P's deceleration from birth onwards is too slow; he continues growing at too high a speed. That is, that the chance of P's deceleration from the $(t_2 - t_1)$ period to the $(t_3 - t_2)$ period belonging to the group G's deceleration over the same period is very small.

We must be careful at the outset to get the exact sense of these hypotheses, to distinguish between populations and samples. The standardising data are all samples. The various extant data sample each somewhat different groups of children, or, in the statistician's

terminology, somewhat different populations. The Brush Foundation data sample the population of healthy white Cleveland children of North American ancestry and above-average economic status. The Fels data sample a somewhat similar population, though perhaps not quite so economically restricted; the Iowa data sample is again from a similar population, though with a different regional restriction. The Harvard School of Public Health group samples a considerably lower economic level, and so on. The data of each of these studies may be taken for use in a particular problem, and it remains to be shown whether in fact some of these samples differ sufficiently for us to consider that they come from truly different populations, or could be obtained by repeated sampling of the same one. In the statistical sense these samples are undoubtedly very poor ones. The children included in the sample, or Growth Study, were not selected either at random or by some other fixed procedure from the population which they are a part of; they were included for accidental reasons, such as willingness to co-operate or being personally known to the investigator. Doubtless the estimation of population figures from such samples is biased, but we can do little but note the fact and pass on. The administrative problems of a Growth Study are formidable enough already; nevertheless if the results are to be used as standards for child growth in the region, it is desirable for new studies to adopt as nearly as possible some recognised sampling procedure, and for the older ones to adjust themselves to some set pattern by controlling their future admissions.

In testing hypotheses A, B or C, then, we are concerned with whether child P is to be regarded as a member of the population or group of which our Growth Study data is an imperfect sample. Thus we must assess our child P's figure against the mean and variability of the population, and these two must be estimated from the mean and variability of our sample data.

In testing hypothesis A, in the case of stature, for example, we proceed as follows: Presuming the distribution of stature in the population of children is normal or very nearly so, the value for P is subtracted from the estimated population mean for the group of children (say white, North American, above-average economic status, healthy) to which he should on the face of things belong. The resulting difference is then compared to the estimated standard deviation of the population. If the hypothesis is that P is *abnormal* (either too tall or too short) and one wishes to misclassify only 5 children out of 100, then P is said

to be abnormal if his difference from the mean is twice (or more accurately 1.96 times) the standard deviation. If the hypothesis is that P is too *short* and if one wishes to classify as pathologically short (and so probably maltreat medically) only 5 normal children in 100, then the level of decision is at a difference of 1.65 times the standard deviation. It should be remembered, however, that this second, one-sided test is far more sensitive than the two-sided test to non-normality in the population distribution. Thus for such measurements as waist or hip circumference, which seem mildly skewed, the one-sided test is suspect, though the two-sided probably suitable. For more heavily skewed distributions, such as weight, even the two-sided test becomes questionable, and quite inapplicable in such very highly skewed measurements as subcutaneous tissue thickness by caliper measurements on tissue folds. There are various possible approaches to these skewed measurements; perhaps the simplest would be by finding transformations, such as log weight instead of weight, for which distributions will approximate normality. Little practical work has yet been done on this, and it will take considerable experience of transformed variates before we can be sure a particular transformation will be in all cases suitable for a particular measurement. Alternatively, statisticians are seeking some form of the test which is less sensitive to skewness (see, for example, reference 32), though less efficient (that is, discovers fewer pathological cases for a given size of standardising sample). Comparison by means of percentiles is often resorted to; the trouble here is that this procedure is less efficient in the statistical sense than would be the *t*-test equivalent on a transformed variate. The location of the population 95th percentile from a sample is done less accurately than the location of the point representing twice the standard deviation; so unless the number of standardising cases is very large, the point taken as 95th percentile for the population may in reality be nearer the 90th or 100th.

This example is taken from the pediatrician's viewpoint; the anthropologist may often be interested in testing the mean figures for one whole sample against the mean of another sample or against a population mean estimated from another series of samples; for this second situation hypotheses A, B and C read equivalently as in the first.

Hypothesis A

The hypothesis A is answered entirely by cross-sectional data; if one is only interested in the truth or falsity of A, the collection of a

longitudinal series would be a ridiculous expense of time and money. This remains true even though a series of independent samples for each year (the pure cross-sectional technique) is a rather less statistically efficient way of estimating the population means at each year than is a series of mixed longitudinal data, or, as Yates calls them, partial replacement samples. By efficiency in this connection is meant the number of subjects needed to give a particular standard error to the population mean—in other words, to estimate the population mean with a particular degree of accuracy. Yates (31, p. 260) finds that fewer subjects would be needed if a mixed longitudinal series were available, but not very many fewer. Consequently the cost of prolonging the study over two or more years and locating a proportion of the same children twice would far outweigh the advantage.

Estimates of population mean from mixed longitudinal series

If, however, a mixed longitudinal series *is* available anyway, collected for the purposes of hypotheses B and C, a somewhat improved estimate of the population mean at each year can be obtained; an estimate improved, that is, over that obtained by treating each year's figures as though purely cross-sectional or independent. Wilson remarked this, and Yates, in a definitive monograph on sampling methods published shortly after the first draft of the present paper was circulated, carries its consideration a good deal further and with considerably greater clarity of exposition (31, see particularly pages 45, 175-182, 233-235, 260-262). The results given in Yates' book were in the main obtained by Patterson who has recently extended them considerably (34).

Notation

The problem of a suitable notation is not altogether simple. We have endeavored to devise one which will meet the requirements of precision and of reasonable simplicity in typesetting, at the cost of somewhat unfamiliar symbolism:

N_h = total number of subjects at time h,

m_h = mean of all subjects at time h,

v_h = variance of all subjects at time h,

$s_h = \sqrt{v_h}$ = standard deviation of all subjects at time h.

The next group of symbols relates to subjects present at time h who were also present at time $(h - 1)$:

L_h = number of such subjects,

$\bar{l}_h$ = mean of these subjects at time h,

$V(l_h)$ = variance of these subjects.

For subjects present at time h but not present at time $(h-1)$ we use

C_h, $\bar{c}_h$, $V(c_h)$ = number, mean, and variance.

We shall also need to consider subjects present at time h according to their presence or absence at time $(h+1)$. These we denote by primed symbols,

L'_h, $\bar{l}'_h$, $V(l'_h)$, and C'_h, $\bar{c}'_h$, and $V(c'_h)$, respectively.

Note that

$$N_h = L_h + C_h = L'_h + C'_h.$$

We shall also want

$\mathrm{cov}(l'_{h-1}, l_h)$ = covariance of measurement of subjects present at both times h and $(h-1)$.

We remark that variances and covariances are calculated with divisor $n-1$ rather than n.

Let us designate by λ the percentage longitudinality from t_1 to t_2, that is, the proportion of children present at time t_2 who were also present a year earlier at t_1. Let μ be the percentage cross-sectionality, or proportion of children present at t_2 and not present at t_1: $\mu = 1-\lambda$. Now if we have only a purely cross-sectional sample at t_2 from which to estimate the population mean for that age, the best estimate actually equals the sample mean. But if we have some members at t_2 previously present at t_1, there are two ways in which the t_2 population mean can be estimated. It can firstly be estimated by reference to the longitudinal element only. This is done by taking the longitudinal element sample value at t_1 and adding to it the mean increase from t_1 to t_2 obtained by the regression of the longitudinal element t_2 values on their corresponding t_1 values. The population mean at t_2 can secondly be estimated from the new values at t_2, the cross-sectional element not present at t_1. This is done quite directly, as if the longitudinal element did not exist. These two estimates of the t_2 mean are independent, and the statistically single best estimate of the t_2 mean is obtained by

weighting and combining the two. Yates gives the expression for this estimate of t_2 mean as

$$M_2 = \frac{L_2\{\bar{l}_2 + b(m_1 - \bar{l}'_1)\} + C_2(1-\mu r^2)\{\bar{c}_2\}}{L_2 + C_2(1-\mu r^2)}, \qquad \text{A}$$

where b is the regression coefficient of measurements at t_2 on measurements at t_1, and r is the correlation coefficient between these measurements. Yates remarks that if the standard deviations on both t_1 and t_2 occasions are nearly equal (which is sometimes but not always the case in growth data) it is better to replace b in equation A by r, since r is less subject to errors of estimation. The fact that expression A is the weighted combination of the two possible estimates is not immediately apparent when it is presented in this form, but it may be noted that the part inside the first curly bracket represents the regression estimate, and the single term in the second curly bracket $\bar{c}_2$, the estimate from the added cross-sectional t_2 values.

Equation A, then, gives us the best estimate of the population mean value on the second of two occasions given a mixed longitudinal series covering two years only. It should be used in testing hypothesis A under these conditions.

Such a situation is unlikely to occur in growth studies; it is too restricted. As a rule there are not two occasions but nearer ten or more, and if we have a mixed longitudinal series given us, one would like, in estimating the population mean at t_6, say, to take into account not only the values at t_5, but also the information provided by the values at t_1, t_2, t_3, and t_4. This information can be worked in by combining numbers of independent estimates as for the simple 2-year case above, but is best presented in a slightly different form, as a recurrence equation relating the mean at t_2 to the mean at t_1, the mean at t_3 to the mean at t_2, and so on. Patterson (34) gives this expression, relating M_h, the best estimate from all the previous years of the population mean at year t_h, to M_{h-1}, the estimate of the population mean of the preceding year t_{h-1}, as

$$M_h = (1-\phi_h)\{\bar{l}_h + \frac{rs_h}{s_{h-1}}(M_{h-1} - \bar{l}'_{h-1})\} + \phi_h \bar{c}_h, \qquad \text{B}$$

where ϕ_h is defined below, and where r is an overall estimate obtained from all the groups of $(h, h-1)$ values in the data, and is calculated as

$$r = \frac{S_h \Sigma\{(l'_{h-1} - \bar{l}'_{h-1})(l_h - \bar{l}_h)\}}{\sqrt{S_h \Sigma(l'_{h-1} - \bar{l}'_{h-1})^2}\sqrt{S\Sigma(l_h - \bar{l}_h)^2}} .$$

Where S_h indicates that the totals of the sums over all pairs of years are to be taken. (Details of computation are in the appendix.) This method of calculating r is adopted because r is, of course, merely an estimate of ρ, the population correlation coefficient, and *ex hypothesi* for the formulae of this section to be valid ρ does not vary from one pair of years to the next (see below); thus each set of paired $t_{h,h+1}$ values furnishes an estimate of ρ, and the more used in the estimation, the better.

The value of ϕ_h is given by a recurrence relation in terms of ϕ_{h-1}:

$$1 - \phi_h = \frac{L_h}{N_h - r^2 C_h\left(1 - \frac{\phi_{h-1}}{C_{h-1}} L_h\right)}, \qquad C_{h-1} \neq 0. \qquad \text{C}$$

In the special case where the same fraction of units is replaced each year and the number of units is constant, we have a somewhat simpler expression

$$1 - \phi_h = \frac{\lambda}{1 - r^2(\mu - \lambda\phi_{h-1})} . \qquad \text{D}$$

Under these conditions ϕ_h approaches, for increasing h, the limiting value

$$\phi_{lim} = \frac{-(1 - r^2) + \sqrt{1 - r^2}\sqrt{1 - r^2(1 - 4\lambda\mu)}}{2\lambda r^2}. \qquad \text{E}$$

For any h greater than 2 the use of equation E leads to little loss of accuracy in estimating M_h; but only for $h > 3$ can equation E be safely used in estimating variance. In any case the assumption of a constant replacement fraction is rarely realised in growth studies, so that we are usually forced to use equation C.

Equation C gives, it may be noted,

$$\phi_h = 0, \text{ if } C_h = 0, \text{ and}$$
$$\phi_h = 1, \text{ if } L_h = 0.$$

In the case where $C_{h-1} = 0$, equation C is indeterminate (since ϕ_{h-1} is then also zero). In such a case we are forced to use the expressions, given in Patterson's paper:

$$1 - \phi_h = \frac{L_h}{N_h - r^2 C_h \{1 - L_h V(M_{h-1})/v_{h-1}\}}, \qquad \text{C'}$$

$$\frac{V(M_{h-1})}{v_{h-1}} = \frac{1 - r^2}{N_{h-1}} + r^2 \frac{V(M_{h-2})}{v_{h-2}}, \qquad C_{h-1} = 0, \qquad \text{C''}$$

$$\frac{V(M_{h-2})}{v_{h-2}} = \frac{\phi_{h-2}}{C_{h-2}}, \qquad C_{h-2} \neq 0. \qquad \text{C'''}$$

Here $V(M_{h-1})$ is the variance of M_{h-1}, and can be calculated provided C_{h-2} is not zero. If $C_{h-2} = 0$, we can take the recurrence relation farther back, until we find some C_{h-k} not zero. An illustration is provided in the appendix. Notice that if C_h is zero as well as C_{h-1}, then $\phi_h = 1$, and no difficulty arises (see Appendix, Example 1). If a year occurs at which there are no subjects at all, i. e., $N_h = 0$, then for the following year ϕ_h is calculated from equation C, using C_{h-2} and ϕ_{h-2} for C_{h-1} and ϕ_{h-1}, and substituting r^4 for r^2.

All these formulae B, C, D and E assume one thing which must be made quite explicit. They are strictly true only if the correlation coefficients of the longitudinal elements remain constant for each adjacent pair of years, i. e. r for the t_h and t_{h+1} values equals r of the t_{h+1}, t_{h+2} values and so on. Secondly, the assumption is made that the correlation between values two years apart is r^2, three years apart r^3, and so on,—that is, that the correlation of t_h and t_{h+2} values equals r^2, of the t_h and t_{h+3} values, r^3. These conditions seem to be well approximated in human growth data except perhaps in the first postnatal year and to a smaller extent during adolescence. In addition Patterson has shown that a fairly small departure from these conditions affects his estimates to a very small degree only. A test of the second of these conditions is shown in Fig. 1. The heavy lines represent the condition when $r_{h,h-2} = r^2_{h,h-1}$, $r_{h,h-3} = r^3_{h,h-1}$ and so on, with the first $r_{h,h-1}$ at the four different levels .98, .97, .96 and .95. The various thin lines represent actual correlations from all the published material available. In the left half of the figure are plotted series for ages before adolescence, t_h being from 9 to 12 (i. e. the highest age being this). In the right half of the figure the ages include adolescence, t_h being 16 or 17. It will be seen that prior to adolescence the data fit the hypothetical conditions quite remarkably well, but during adolescence, as is well known, r decreases and the thin lines depart from the condition given by the thick. So far as the evidence of these data goes, however, the disturbance during adolescence seems to be temporary; once adolescence is over, the thin

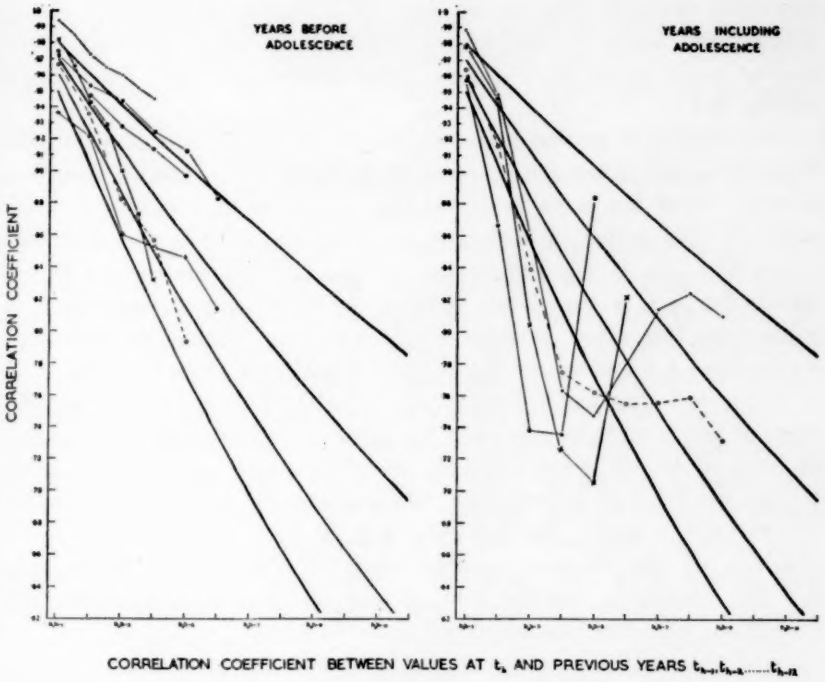
YEARS BEFORE ADOLESCENCE
YEARS INCLUDING ADOLESCENCE
CORRELATION COEFFICIENT
CORRELATION COEFFICIENT BETWEEN VALUES AT t_h AND PREVIOUS YEARS $t_{h-1}, t_{h-2}, \ldots t_{h-n}$

FIG. 1. TEST OF CONDITION $r_{h,h-k} = r^k_{h,h-1}$

Heavy lines represent condition for four values of $r_{h,h-1}$. Light lines actual correlations from various data, for years before adolescence (Fig. 1A) and years including adolescence (Fig. 1B).

Fig. 1A. Years before adolescence:

●——●	35 boys	bi-iliocristal, pure longt.
×——×	20 girls	$h = 9$. Meredith & Carl, 1946
⊙——⊙	Stature	275 girls, Harvard Growth
○---○	Weight	Study, pure longt. $h = 12$
■——■	Boys	Stature. Brush data, mixed
+——+	Girls	longt. $h = 12$. Simmons, 1944

Fig. 1B. Years including adolescence:

●——●	Stature	Harvard Growth Study, 275
○---⊙	Weight	girls, pure longt. $h = 16$. Wilson, 1935
×——×	Girls	Stature, Brush Foundation data,
■——■	Boys	Simmons 1944, mixed longt.
		$h = 16$ girls, 17 boys.

lines seem to get back onto the thick tracks at just about where the condition demands they should be. (The figure, incidentally, demonstrates graphically the well-known fact that prediction of adult stature can be made better from stature before adolescence than from stature during it.)

We have now got an equation for the best estimate of population mean at t_6, say, when the values at t_1, t_2, t_3, t_4 and t_5 are all taken into account. But the same principle can be pushed still further; we can obtain a still better estimate for t_6 if we also take into account the values at t_7, t_8, t_9, t_{10}, and so on. Obviously if we are casting up our results for t_6 at a time when none of our children have reached t_7 we cannot use this extra information; it is still to come. But very often we do have a lot of values for t_7 and thereafter. If our growth study is of the sort that has a distinct end as all its subjects reach maturity more or less simultaneously as in the Harvard Growth Study, we should, at least in theory, recalculate all our data when the end is reached, taking every bit of information into account for each year's estimate. If our growth study, like the Fels, is built on a continuous intake of children, we should every so often compute completely new estimates for our standards, taking into account all the data available and modifying our previous estimates accordingly.

Unfortunately the expression for the best estimate of the mean of t_h taking into account all the values for years both before and after t_h is computationally complicated, and Patterson confines himself to giving the estimate which includes information from only a single year after t_h in addition to all the years before it. We will denote this estimate by ${}_{h+1}M_h$ to distinguish it from the estimate M_h of equation B which did not take into account the values available at t_{h+1}. The equation giving ${}_{h+1}M_h$ is

$$ {}_{h+1}M_h = M_h - \frac{rs_hC_{h+1}\phi_h}{s_{h+1}C_h}\{M_{h+1} - \bar{c}_{h+1}\}, \qquad C_h \neq 0, \qquad \text{F} $$

where M_h and M_{h+1} are obtained from equation B and its equivalent for M_{h+1}, and ϕ_h from equation C. When $C_h = 0$ equation F no longer holds, and as in the case of ϕ_h some added complexity results. The full expression for ${}_{h+1}M_h$ is

$$ {}_{h+1}M_h = M_h - rs_h\frac{C_{h+1}}{s_{h+1}}\,\frac{V(M_h)}{v_h}\{M_{h+1} - \bar{c}_{h+1}\}. \qquad \text{F}' $$

This reduces to equation F when $C_h \neq 0$, but otherwise must be used

directly, with a value for $V(M_h)/v_h$ found from equations C″ and C‴ with the appropriate changes of subscripts. An illustration is provided in Example 2 in the Appendix. Note however that if $C_{h+1} = 0$, then ${}_{h+1}M_h = M_h$ whether or not $C_h = 0$, so that if $C_{h+1} = 0$ it is never necessary to use equation F′.

Relative efficiency of estimates of yearly mean size and mean velocity of growth

We have now seen that, given a mixed longitudinal sample of children, there are several ways of estimating the population mean value at each year for a given dimension. We can use the data as though they were purely cross-sectional and take as our population value m_h, the mean of all the measurements at t_h. We can take into account values prior to t_h and take as our estimate M_h from equation B. We can take into account these values and the additional value at t_{h+1} and use ${}_{h+1}M_h$ from equation F. Finally we can in theory take into account an infinite number of after-h readings and estimate the mean as ${}_{\infty}M_h$. It is essential to realise exactly how these different estimates compare with one another. All are *unbiased*, that is to say, as the number in the sample increases, all get nearer and nearer to the true population mean value. But their *efficiencies* differ greatly, and increase in the order given above. That is to say, the number of cases needed to establish the population mean securely within given limits—perhaps 5 per cent of the mean value—will be considerably greater if we take m_h as our estimate than if we take M_h, and M_h in turn will need more cases for its establishment than will ${}_{h+1}M_h$. In other words, we are letting a great deal of information slip through our fingers if we use m_h (and spending a great deal of extra money for a given accuracy). We use more of our available information by calculating M_h, still more by calculating ${}_{h+1}M_h$, and the maximum possible by calculating ${}_{\infty}M_h$. Patterson gives the relative efficiencies of the three estimates in relation to ${}_{\infty}M_h$, for various values of r and two values of λ, and his figures are reproduced in Table 1. In the left-hand half of the table are the relative efficiencies for the estimate of the population mean; for example, at the very usual growth study figures of $r = .95$, $\lambda = .67$, the efficiency of m_h as an estimate is only .33 or 33 per cent of that of ${}_{\infty}M_h$; M_h is 66 per cent efficient, ${}_{h+1}M_h$ is 76 per cent efficient. It will be noticed that r affects the efficiency much more than does λ, but as both get higher so the necessity of using M_h and ${}_{h+1}M_h$ becomes greater. Patterson

has assumed equal replacement fractions, i. e. equal λ, from one year to the next and equal standard deviations at each year in constructing this table, but removal of these restrictions will not affect the figures very much.

TABLE 1

Efficiencies, relative to the most accurate possible estimate, of population yearly means and year-to-year increments from mixed longitudinal series
Taken or computed from Patterson (34).

	ESTIMATES OF POPULATION MEAN AT YEAR h BY			ESTIMATES OF POPULATION MEAN INCREMENT FROM YEAR $h-1$ TO YEAR h BY			
r	m_h	M_h	$_{h-1}M_h$	$m_h - m_{h-1}$	$M_h - M_{h-1}$	$M_{\text{incr}_{1,2}}$	$M_h - {}_{h-1}M_h$
			$\lambda = 1/2$				
0	1.00	1.00	1.00	1.00	1.00	1.00	1.00
.25	0.97	0.98	1.00	0.97	0.98	0.99	0.99
.50	0.87	0.93	0.99	0.85	0.93	0.96	0.97
.60	0.80	0.90	0.98	0.76	0.90	0.94	0.96
.70	0.71	0.86	0.97	0.65	0.86	0.92	0.95
.80	0.60	0.80	0.94	0.50	0.80	0.90	0.94
.90	0.44	0.72	0.87	0.30	0.72	0.90	0.93
.95	0.31	0.66	0.78	0.16	0.66	0.88	0.93
1.00	0.00	0.50	0.50	0.00	0.50	0.00	1.00
			$\lambda = 2/3$				
0	1.00	1.00	1.00	1.00	1.00	1.00	1.00
.25	0.97	0.99	1.00	0.97	0.99	0.99	0.99
.50	0.88	0.94	0.99	0.86	0.94	0.96	0.98
.60	0.82	0.91	0.98	0.79	0.91	0.95	0.97
.70	0.73	0.87	0.96	0.69	0.87	0.94	0.96
.80	0.62	0.81	0.92	0.54	0.81	0.93	0.96
.90	0.46	0.73	0.84	0.33	0.73	0.93	0.95
.95	0.33	0.66	0.76	0.19	0.66	0.90	0.95
1.00	0.00	0.50	0.50	0.00	0.50	0.00	1.00

These efficiencies are important in any situation in which it is necessary to take into account the standard error of the estimate of the mean; the efficiencies are, of course, inversely proportional to the square of the standard error, the variance. Thus to say that one estimate has a higher efficiency than another is to say its standard error is lower. The size of the standard error is important in comparing different samples of children one with another, in fitting curves to the yearly mean values, in testing hypothesis A, and in estimating the

velocity of growth from one year to another,—indeed in practically every situation likely to arise in work on growth. The variances of these various estimates are given by Patterson as

$$V(m_h) = v_h/N_h, \qquad \text{G}$$

and for $C_h \neq 0$,

$$V(M_h) = v_h\phi_h/C_h. \qquad \text{H}$$

When $C_h = 0$ we have

$$V(M_h) = v_h \left\{ \frac{1-r^2}{N_h} + \frac{r^2 V(M_{h-1})}{v_{h-1}} \right\}, \qquad \text{H}'$$

which can be prolonged backward until some time t_k is found when $C_k \neq 0$ and $V(M_k)$ can be estimated by equation H. Note that equation H′ holds *only* when $C_h = 0$. Next

$$V({}_{h+1}M_h) = \frac{v_h\phi_h}{C_h} \left\{ 1 - \frac{r^2 C_{h+1}\phi_h(1-\phi_{h+1})}{C_h} \right\}, \qquad \text{I}$$

which is the most convenient form for computation but holds only when $C_h \neq 0$. The general form, which holds for $C_h = 0$, and reduces to equation I when $C_h \neq 0$, is

$$V({}_{h+1}M_h) = V(M_h) \left\{ 1 - r^2 C_{h+1}(1-\phi_{h+1}) \frac{V(M_h)}{v_h} \right\}. \qquad \text{I}'$$

When $C_h = 0$, equation I′ must be used, and $V(M_h)$ computed from equation H′, or repeated applications of it if necessary. Lastly,

$$V({}_{\infty}M_h) = \frac{s^2\phi}{N(2\mu-\phi)}, \qquad \text{J}$$

expression J applying only to circumstances of equal replacement of units on each occasion, that is, equal μ for each pair of years.

In Table 1 are given in the right-hand half the relative efficiencies of four estimates of increment, also from Patterson. In presenting this now we are anticipating somewhat, and we will return to its consideration again later, but it may be as well to point out directly that population estimates of *rate of growth* can best be obtained from two adjacent estimates of mean yearly values as given above. Which estimate of mean yearly values is taken naturally affects the efficiency with which the increment is assessed, and since it is really increments we are most interested in, these figures are of prime importance and are perhaps our best practical guide as to whether it is really necessary

to compute ${}_{h+1}M_h$ instead of M_h, etc. The four estimates of mean increment are $m_h - m_{h-1}$ (difference of overall means, data treated completely cross-sectionally), $M_h - M_{h-1}$, $M_{\text{incr}\,1,2}$ which is the best estimate [3] available using only the present and previous occasion figures, and $M_h - {}_hM_{h-1}$. The efficiency of the first of these estimates is hopelessly low, of $M_h - M_{h-1}$ none too good, but of $M_h - {}_hM_{h-1}$ very excellent. This is in terms as before of the efficiency of an estimate using all possible information, i. e. ${}_\infty M_h - {}_\infty M_{h-1}$. The estimate $M_h - {}_hM_{h-1}$ is all we need ever compute in any foreseeable circumstances, so that in computing increments we may need to invoke equations B and F, B for the second of the two years and F for the first of the two. (Note that ${}_hM_{h-1}$ is the same form as ${}_{h+1}M_h$ but for the previous year.)

Testing hypothesis A

In testing hypothesis A we need estimates of the population mean at a given year t_h and of the population standard deviation at the same year. The way to estimate the mean if the series is mixed longitudinal is as outlined above, by means of ${}_{h+1}M_h$ if t_{h+1} data are available, by M_h if they are not. The standard deviation is obtained quite straightforwardly from the values at t_h alone, estimated as s_h.

Hypothesis B

Hypothesis B demands longitudinal data, if the test of it is not to be hopelessly insensitive. Either pure or mixed longitudinal data will serve, but the mixed only if used as a longitudinal series, and not as a cross-sectional. The simplest way to do this is to calculate each individual increment from t_1 to t_2 and then average them for the mean, and subsequently estimate the population standard deviation from them directly, as has been done in Dr. Thompson's data. The increment from t_1 to t_2 of the child P is then compared to the estimated mean increment of the population from which P is presumed drawn, and the difference between the two figures compared to the estimated standard deviation of the increments.[4]

[3] Discussed below; see equation N.

[4] This presumes a normal or near-normal distribution of increments. Though this point is one of very considerable importance in several respects, the only data on it I can find are Boas' (28) on height at adolescence, which seem reasonably normal, and Palmer's (4) monthly increments of weight, which are somewhat skewed. It is most desirable that frequency distributions of annual increments should be published.

This way of dealing with a mixed longitudinal series is equivalent to picking out from it the pure longitudinal data from children present at both t_1 and t_2 and dropping the added cross-sectional data, that is, those children who appeared at t_1 but not at t_2 and those who appeared at t_2 and not t_1. Only λ per cent of the data are used. This is not in fact the most efficient way to estimate the mean increment, since the dropped cross-sectional data can contribute a little, though not usually much, more information. It is however a very convenient way of estimating mean increment in many circumstances, since the individual increments have to be tabulated to get the standard deviation of increment for use in testing hypothesis B, whatever way of estimating the mean increment is used. How the best estimate of mean increment, using all the information, can be obtained is discussed below; but the added information over that provided by the longitudinal elements alone is very often not worth the extra labour of computing and may provide scarcely any greater accuracy unless the percentage longitudinality of the data is improbably low—perhaps about 40 per cent.

The uselessness of reporting overall means and standard deviations only

First, however, a different point. The vast majority of growth study data up to the present have been published only as overall yearly means and standard deviations. Each overall mean is a mixture of its λ longitudinal elements and its μ cross-sectional ones. But nobody is to know how much of the data is longitudinal and how much cross-sectional, since only the total numbers at each year are given. The invitation is plain to treat the data as if simply cross-sectional, and indeed many extensive studies of so-called rate of growth do precisely this (e. g. 33). Now what will happen to the test of our hypothesis B if the mixed longitudinal data are treated as though they were cross-sectional?

This is the point already examined by Shuttleworth (1) and Wilson (2). We want to know what will happen to our estimates both of population mean increment, M_{incr}, and of population standard deviation of increment s_{incr}. Let us deal with s_{incr} first. Suppose we have estimated poulation standard deviations s_1 and s_2 at time t_1 and t_2, and consider initially a pure longitudinal series. There are two ways of working out s_{incr}, both algebraically equivalent and providing exactly the same answer. The increments can be individually tabulated and their standard deviation worked out; this is the straightforward and obvious way, but requires

access to the original raw data. Alternatively, we can calculate s_{incr} if we know s_1 and s_2 and r, the coefficient of correlation between the values at t_1 and their corresponding values at t_2 by the expression

$$s_{incr} = \sqrt{s_1^2 + s_2^2 - 2rs_1s_2}. \qquad \text{K}$$

This second method of calculating s_{incr} would have to be used if the investigator knew only the means, sigmas and correlation coefficient, and had no access to the raw data.

Now the crux of the matter is this: in a pure longitudinal series r will be of the order of .95, while in a pure cross-sectional it will be 0. This is stretching the terminology of course, and it is perhaps better to put it as follows: a mixed series consists of a longitudinal element and a cross-sectional element. Let us call r for the mixed series the correlation coefficient for the λ per cent of children who do have two values, in other words for the longitudinal element only; it will usually be, as we have already said, about .95. There remain the μ per cent of children, who can also be paired one at t_1 against one at t_2 if there are equal total numbers on the two occasions. But in these pairings, made at random, one value of the pair will have no relation to the other value, and thus the correlation for these μ children will be 0, on the average. The whole set of increments will thus be made up of a proportion for which the standard deviation will be given by equation K with $r = .95$, and a proportion μ for which the correlation is 0. In the combined set, with equal numbers on the two occasions, we shall have

$$s_{incr} = \sqrt{s_1^2 + s_2^2 - 2\lambda rs_1s_2}. \qquad \text{L}$$

As the percentage longitudinality rises, so s_{incr} drops. Also, for a given percentage longitudinality, as r increases, s_{incr} drops. Thus if we ignore the last term in equation L and treat our mixed series as though it were cross-sectional with $\lambda = 0$, we grossly over-estimate s_{incr}. Just how grossly can easily be calculated approximately. It can be seen from equation L that when $s_1 = s_2$ this approximate mixed longitudinal estimate of s_{incr} is $100/\sqrt{1 - \lambda r}\%$ of the estimate of s_{incr} got by treating the data as though they were purely cross-sectional, that is, by putting $\lambda = 0$. If we take $r = .95$, we have s_{incr} from L as 38%, 49%, 58%, 72% and 84% of cross-sectional s_{incr} for $\lambda = .9, .8, .7, .5$ and .3 respectively. Even if r reaches its lowest values of about .65, seen only in the first year of life, s_{incr} from equation L would still be 64%, 69%, 74%, 82% and 90%

for the same values of λ. So much for the estimation of s_{incr} by ignoring longitudinality. Is the estimation of M_{incr} equally poor? A glance back to Table 1 suffices to show that it is. In the right-hand block are the efficiencies of various estimates of mean increment relative to the best possible estimate, and the estimate ignoring longitudinality is the left-hand column of the block, $m_h - m_{h-1}$. Its efficiency is pathetically bad; for $\lambda = 2/3$, it is only 54% efficient at $r = .8$, 33% efficient at $r = .9$, and 19% efficient at $r = .95$. The point seems worth driving home that a little biometrical study is worth a lot of collector's instinct; at this level of longitudinality a growth study which pauses to extract all its information from the data will give as accurate an estimate of mean rate of growth as another study having five times as many subjects on its books, but reporting its data in the time-honoured way.

Table 1 gives efficiencies relative to the best possible estimate; we have said above that the estimate from the contained longitudinal element only of a mixed series is often nearly as good as the best estimate and more convenient. How, then, does the overall means estimate $m_h - m_{h-1}$ compare with the estimate $\bar{l}_h - \bar{l}'_{h-1}$ obtained solely from the longitudinal element? Consider first the isolated longitudinal element with means $\bar{l}_h - \bar{l}'_{h-1}$, and then add cross-sectional cases at t_h and t_{h-1}. We are adding in subjects that were otherwise unavailable and in general increasing the number in the sample leads to a better estimate of the population figure. Against this in the present case, however, we are losing accuracy by introducing a cross-sectional increment with a far larger standard deviation than the longitudinal increment, and yet treating the two elements as equally important in relation to their numbers by using the inefficient overall means estimate. This second effect overrides the effect of the total numbers being greater until such time as the added cross-sectional numbers are increased very considerably over the initial longitudinal ones. A time does come however when our overall means estimate becomes as good as the original estimate founded on fewer cases, but purely longitudinal ones.

Wilson (2) worked out the mathematics of this and found that the time came when the ratio

$$\frac{\text{Children present at } t_1 \text{ but absent at } t_2}{\text{Children present at } t_1 \text{ and at } t_2} \geqq \frac{2r-1}{1-r}$$

given the simplifying assumptions of equal total numbers and equal

standard deviations at t_1 and t_2. In the terminology used above this condition is

$$\frac{\mu}{\lambda} \geqq \frac{2r-1}{1-r},$$

i. e.

$$\lambda \leqq \frac{1-r}{r}. \qquad \text{M}$$

I have put down in Table 2 the values for λ corresponding to various values of r which give equality in expression M. These values for λ are those below which the percentage longitudinality of a mixed series has to fall before treating the values cross-sectionally gives as good an

TABLE 2

Values of minimal percentage longitudinality λ in a mixed series for which estimation of mean increment is better from longitudinal values only than by treating overall values cross-sectionally

λ given by $\lambda = (1-r)/r$ where r is correlation coefficient of longitudinal element.

r	λ
.98	.02
.96	.04
.94	.06
.92	.09
.90	.11
.85	.18
.80	.25
.75	.33
.70	.43
.65	.54
.60	.67

estimate of mean increment as does dropping the cross-sectional element completely. Suppose a given series has a longitudinal element whose correlation is .90; by reference to the table we find that dropping the cross-sectional element is more efficient unless the series is under 11 per cent longitudinal, a most unusual occurrence. Only when the correlation gets very low does the longitudinality have to be relatively high before cross-sectional reporting becomes the more efficient of the two methods.

An example will show just how badly the overall cross-sectional estimation of mean increments and standard deviations serves us. We

will test our hypothesis B first in the incorrect cross-sectional manner and then taking our estimations from the pure longitudinal element only. The Brush Foundation data are a convenient source. In the monograph devoted to reporting the data (5) only overall means and standard deviations for each year are given and we cannot disentangle cross-sectional and longitudinal elements. But in a previous publication dealing with heights and weights only (6) this usual method is supplemented by tables of actual yearly increments for each t_{h-1} to t_h longitudinal element, correlation coefficients for the longitudinal elements and tables of percentage longitudinality. (This is nearly, but still not quite, all that is needed to make the best possible test of hypothesis B; the authors omit to give the means for each year for the longitudinal and cross-sectional elements separately, i. e. $\bar{l}_h$ and $\bar{c}_h$ and these also are required.) We will take the male stature increment from 6 to 7 years. The overall mean tables show there were 113 boys at age 6 with mean 117.24 cms. and 104 boys at age 7 with mean 123.89 cms. The mean increment by subtracting one of these figures from the other would thus be 6.65 cms. The 84 boys who constitute the longitudinal element however had a mean increment of 6.41 cms. and since $r = 0.97$ and $\lambda = 0.74$, this is a much better estimate of the mean rate of growth from 6 to 7 of the population of boys sampled in this study. The best estimate, from equation F, cannot be made.

Fig. 2 illustrates this difference between the overall $m_h - m_{h-1}$ estimates and the longitudinal $\bar{l}_h - \bar{l}'_{h-1}$ estimates, over the whole growth span, for stature and weight in boys and girls. The solid line represents the longitudinal element estimate $\bar{l}_h - \bar{l}'_{h-1}$ and the dotted line the overall cross-sectional estimate. I have labelled the curves "true" and "false"; both are unbiased, but the "true" curve is much more precise, that is, its points are less subject to sampling error. The greater smoothness of the "true" increment curves is apparent in each graph, and in the top two figures I have indicated the $\pm$ 2 standard errors of the actual mean increments at age 11-12. The "false" increment figure falls outside even these limits. One can without difficulty imagine a pre-adolescent slowing of weight velocity being postulated from the dips in the "false" increment weight curves at 9-12 and 9-10, and a 7-9 spurt in female stature appearing from the lower left-hand graph. The "false" curves obscure the true course of events by their high standard errors, and may mislead the unwary into unwarranted conclusions.

Now as to the estimation of s_{incr}. Presented only with overall yearly

means, we would be in something of a quandary what to use as s_{incr}, but suppose we carried on the same way as for means and calculated s_{incr} as for cross-sectional data by the usual method as $\sqrt{s_1^2 + s_2^2}$. This figure for the Brush male stature data is 6.56 cms. Actually, however,

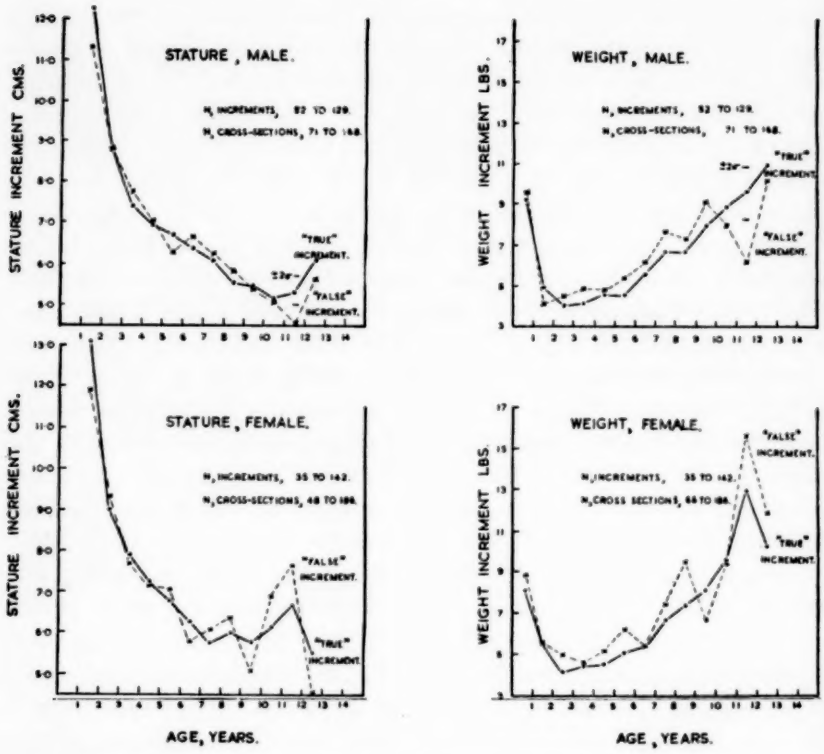

FIG. 2. "TRUE" MEAN INCREMENTS COMPUTED FROM LONGITUDINAL ELEMENT ONLY AND "FALSE" MEAN INCREMENTS COMPUTED FROM OVERALL CROSS-SECTIONAL MEANS IN A MIXED LONGITUDINAL SERIES. BRUSH FOUNDATION DATA, SIMMONS AND TODD, 1938.

Heavy line, "true" increment; dotted line, "false" increment.

the s_{incr} is far less than this, r being .97. From the longitudinal element of 84 boys the estimate of s_{incr} is 1.29 cms., about a fifth of the previous, wrongly-made estimate. Suppose a physician was presented with a boy believed drawn from the same population as the Brush Foundation sample, who had grown just 2.0 cms. from his sixth to his seventh

birthday. If he knew the mean and standard deviation of the 6 to 7 year increment as estimated from the longitudinal element of the study, he would calculate the boy was $(2.0 - 6.41)/1.29 = -1.93$ standard deviation units down from the mean of the population growth. If he took the hypotheses of the child being abnormal, either small or large, he would be very suspicious of abnormality, since the conventional test would demand a figure of ± 1.96; if he took the hypothesis of the child being too short, he would be practically certain he was confronted with an abnormality since the test demands a value of only -1.65, and he would proceed accordingly. But if he did not have these estimates to go by, but only those from the overall cross-sectional figures (which is all in practice he has got for almost any standards, such is the situation at present) he would calculate the boy was $(2.0 - 6.65)/6.56 = -0.71$ standard deviation units below the mean and (unless gifted with common sense) would pay no further attention to the matter. Had equation K been used with $r = .97$ as though the whole series was longitudinal, s_{incr} would have been 1.27 cm., or using equation L with $1 = \lambda$, s_{incr} would have been 3.51 cm.

In short, the inefficiency of our method of reporting and using the standards has completely vitiated the purpose of their collection. It is clear that data reported in the overall mean and standard deviation manner cannot be used as they stand either by the pediatrician for comparing a given child's increment with that of a group, or by the biologist interested in rate of growth.

The best estimate of mean increment

We have already pointed out in passing that the estimate of mean increment from the pure longitudinal element alone, though better than that from overall means, is not the best possible estimate; it wastes the admittedly small extra amount of information provided by the thrown-away cross-sectional cases. Just as both longitudinal and cross-sectional elements can independently furnish estimates of the t_h mean (see above) so they can furnish independent estimates of the t_{h-1} to t_h increment. Each estimate is given its correct weighting and the two combined to give the single best estimate. The formula for the increment from t_{h-1} to t_h taking into account *only* t_{h-1} and t_h values (and not those prior to t_{h-1} or after t_h) was given by Wilson; it is the increment equivalent of expression A for the estimate of M_2 above. We will call this estimate of increment $M_{\text{incr}\,1,2}$; it is calculated as

$$M_{\text{incr}\,1,2} = \frac{A(\bar{l}_h - \bar{l}'_{h-1}) + B(\bar{c}_h - \bar{c}'_{h-1})}{A+B} \qquad \text{N}$$

with

$$A = \frac{V(c_h)}{C_h} + \frac{V(c'_{h-1})}{C'_{h-1}},$$

$$\text{B} = \frac{V(l_h) + V(l'_{h-1}) - 2\,\text{cov}(l_h, l'_{h-1})}{L_h}.$$

The terms A and B are weights attached respectively to the longitudinal element increment and the cross-sectional increment. Under the simplifying conditions of equal numbers and equal variances on both occasions, expression N reduces to

$$M_{\text{incr}\,1,2} = \frac{\lambda}{1-\mu r}\{\bar{l}_h - \bar{l}'_{h-1}\} + \frac{\mu(1-r)}{1-\mu r}\{\bar{c}_h - \bar{c}'_{h-1}\} \qquad \text{N}'$$

where the relative weightings are perhaps more clearly seen.

Equation N is the appropriate expression for computing mean increment when there are only two years in the series; there are other circumstances, to be discussed in a minute, when it may also be very appropriate. But, what we need now, at least theoretically, is a formula for the t_{h-1} to t_h increment which takes into account all the increments before t_h, as did the formula B for M_h, and one which takes into account all increments after t_h also, comparable to ${}_{\infty}M_h$. Such expressions are best obtained simply by estimating the means at t_{h-1} and t_h as detailed in the previous section, and subtracting one mean from another to get the estimate of mean increment. Thus the best possible estimate of increment from t_{h-1} to t_h, taking into account all data before and after t_h would be ${}_{\infty}M_h - {}_{\infty}M_{h-1}$. Calculation of this is excessively laborious, and unnecessary because of the very high efficiency of the estimate $M_h - {}_hM_{h-1}$. This is shown in the right-hand column of the right-hand block of Table 1, from Patterson. These figures hold strictly only for the conditions of equal replacement fraction (equal λ throughout) and equal standard deviation at each year, but they will not be greatly altered by removal of these restrictions. It will be seen (perhaps with relief) that even for high values of r and low values of λ this $M_h - {}_hM_{h-1}$ estimate is over 90 per cent efficient, and there seems no reason to use any more complicated estimate in any foreseeable circumstances in human growth data. Though this estimate can be expressed in terms of the M_h series only, it is computationally simpler to work out the ${}_hM_{h-1}$

series also by equation F, and obtain $M_h - {}_hM_{h-1}$ by simple subtraction; and in any case it is desirable to have the extra 10 per cent or so efficiency provided by ${}_hM_{h-1}$ in the estimate of mean yearly values.

The other three columns in the right-hand block of Table 1 give the efficiencies of other estimates of mean increment. The overall means increment $m_h - m_{h-1}$ we have already discussed. The estimate $M_h - M_{h-1}$ is rather surprisingly a good deal less efficient for commonly met values of λ and r than $M_h - {}_hM_{h-1}$ and so the latter is greatly to be preferred. On the other hand the estimate from the present and previous occasions only, $M_{\text{incr}\,1,2}$, is surprisingly good, and though not quite as efficient as $M_h - {}_hM_{h-1}$ does approach it fairly closely. Most growth studies have most of their λ fractions above $\frac{1}{2}$ and their correlation coefficients above .95, and in such circumstances $M_{\text{incr}\,1,2}$, is not greatly under 90 per cent efficient.

The variances of $M_{\text{incr}\,1,2}$ and of $M_h - {}_hM_{h-1}$ (which would be needed in a comparison, for example, of boys' rate of growth with girls') have been given by Wilson and by Patterson respectively.

$$V(M_{\text{incr}\,1,2}) = \frac{AB}{A+B}, \qquad \text{O}$$

where A and B have been computed in evaluating $M_{\text{incr}\,1,2}$;

$$V(M_h - {}_hM_{h-1}) = V(M_h) + V({}_hM_{h-1}) - 2\,\text{cov}(M_h, {}_hM_{h-1}) \qquad \text{P}$$

$$= V(M_h) + V({}_hM_{h-1}) - 2r\frac{s_h}{s_{h-1}}(1-\phi_h)V(M_{h-1}). \qquad \text{P}'$$

Expression P′ is the most convenient if the variances of the ${}_hM_{h-1}$ series have been calculated. It holds also for $C_{h-1} = 0$, when $V({}_hM_{h-1})$ must be obtained from equation I and $V(M_{h-1})$ from equation H′. When $C_{h-1} \neq 0$ and the variances of the M_h series only have been calculated, the most convenient form is

$$V(M_h - {}_hM_{h-1}) = V(M_h) + V(M_{h-1})$$
$$\times \{1 - r^2\frac{C_h}{C_{h-1}}\phi_{h-1}(1-\phi_h) - 2r\frac{s_h}{s_{h-1}}(1-\phi_h)\} \qquad \text{P}''$$

or directly from the basic statistics by

$$V(M_h - {}_hM_{h-1}) = \frac{\phi_h v_h}{C_h} + \frac{\phi_{h-1}v_{h-1}}{C_{h-1}}$$
$$\times \{1 - r^2\frac{C_h}{C_{h-1}}\phi_{h-1}(1-\phi_h) - 2rs_hs_{h-1}\frac{\phi_{h-1}}{C_{h-1}}(1-\phi_h)\}. \qquad \text{P}'''$$

It would be very desirable to put in the mean increment curves estimated by $M_h - {}_hM_{h-1}$ and $M_{\text{incr}\,1,2}$ in Fig. 2, but unfortunately the absence of figures for $\bar{l}_h$ from the paper cited makes this impossible. The considerations below, however, indicate that such curves would coincide very closely with the pure longitudinal ones.

Estimation of increment from longitudinal element only

This fact introduces the last important point in this section. The addition of the cross-sectional data estimate gives so little extra information under most circumstances as to be scarcely worth the

TABLE 3

Variance of estimate of mean t_{h-1} to t_h increment from pure longitudinal element of mixed series only, in percentage of variance of best estimate, by equation N′, from expression $100(1-\mu r)/\lambda$

λ	$r = .95$	$r = .90$	$r = .80$	$r = .70$	$r = .60$
.95	100.0	100.4	101.0	101.7	102.2
.9	100.4	101.3	102.2	103.5	104.4
.8	101.0	102.4	105.0	107.5	110.1
.7	102.0	104.2	108.6	112.6	117.0
.6	103.3	106.9	113.4	119.8	126.8
.5	105.0	110.3	119.8	130.0	139.9
.4	107.1	114.9	130.0	145.0	161.8
.3	111.8	123.2	146.9	170.1	193.5
.2	119.8	139.9	180.1	219.9	261.2
.1	145.0	210.6	279.9	370.2	460.1

trouble of computing. This can be seen most easily in the somewhat restricted circumstances represented by equation N′. The relative weights of longitudinal and cross-sectional estimates in this t_1 to t_2 case are $\lambda/1-\mu r$ and $\mu(1-r)/1-\mu r$. For $r = 0.95$ these weights work out at longitudinal to cross-sectional 199 to 1 for $\lambda = .9$, 82 to 1 for $\lambda = .8$, 30 to 1 when $\lambda = .6$. A more compact way of recording the same thing is shown in Table 3. The variance of the estimate $M_{\text{incr}\,1,2}$ of equation N′, under the conditions for which the estimate is fully efficient, is

$$V(M_{\text{incr}\,1,2}) = \frac{2s^2(1-r)}{N} \cdot \frac{1}{1-\mu r} \qquad \text{Q}$$

(which can be obtained from equation P by imposing the restriction of

equal numbers and equal variances). The variance of the pure longitudinal series contained within this mixed one would be $2s^2(1-r)/N\lambda$ if we ignore the small correction in s^2 imposed by the fewer numbers, and is thus very nearly $(1-\mu r)/\lambda$ times the variance of the best estimate given in expression Q. For any usual value of r, this fraction is very near unity; it is set down for varying values of r and λ in Table 3. The variance of the estimate of mean increment obtained, for example, by dropping the cross-sectional element of a series with $r = .95$ and $\lambda = .80$ is just 101.0 per cent of the variance of the best estimate of mean increment. The gain in efficiency by using the best estimate is quite trivial for values of λ down to 0.4 or so when r is .90 or above, and down to $\lambda = 0.7$ for $r = .70$ even.

Admittedly this test applies only to the equal replacement, equal standard deviation conditions and to the estimates of increment from the present (t_h) and the previous (t_{h-1}) occasion only. We have already seen, however, that under these conditions the present-previous estimate of increment $M_{\text{incr}\,1,2}$ is 90 per cent or more efficient, and the removal of the equal standard deviation equal replacement conditions will not greatly affect this. Thus while Table 3 may overestimate the relative efficiency of the dropping-the-cross-sectional data scheme for the unrestricted case, it does not do so to any great extent—probably not more than 10 per cent under most foreseeable circumstances. In addition to this, both the longitudinal-element-only estimate of increment $\bar{l}_h - \bar{l}'_{h-1}$ and the present-previous estimate of increment $M_{\text{incr}\,1,2}$ have one advantage (as Patterson points out for the latter) over the best estimate. They are not affected by the pattern of the year-to-year correlations specified above (and in Fig. 1) of r constant from one year to the next and of $r_{h,h-2} = r^2_{h,h-1}$, etc. Departures from this condition do occur to some extent in human growth data at the early ages and during adolescence, and they will tend to make the variance given in expression P somewhat too small, and therefore possibly slightly misleading; the effect, though existent, is not a large one. In circumstances where this pattern of correlations does not quite hold, therefore, the longitudinal-element-only increment and the $M_{\text{incr}\,1,2}$ estimate may actually be theoretically as well as computationally preferable to $M_h - {}_hM_{h-1}$.

The variance of the longitudinal-element only increment is given by

$$V(\bar{l}_h - \bar{l}'_{h-1}) = \{V(\bar{l}_h) + V(\bar{l}'_{h-1}) - 2\,\mathrm{cov}(\bar{l}_h, \bar{l}'_{h-1})\}/L_h. \qquad \text{R}$$

At this point the reader may be referred to the two examples in the

appendix, where, in Tables 5 and 7, he will find tabulated the practical tests of these theoretical findings. In particular the estimates of change $\bar{l}_h - \bar{l}'_{h-1}$ and $M_h - {}_hM_{h-1}$ should be compared (they will be found in the last and last-but-one blocks of these tables). The two estimates differ very slightly, as predicted; this is seen better in Example 2, which is more extensive than Example 1, and does not have a large number of $C_h = 0$ as Example 1 does. The average difference is about .02 to .03 cms. for half-yearly gains, representing about 3 per cent of the average half-yearly increment. The difference is on the average about one-third of the standard error of the $M_h - {}_hM_{h-1}$ estimate of mean increment; in only one instance out of the 32 calculated does it exceed one standard error, and then only just. The variances of these two estimates of increment are about the same, as predicted, neither being clearly systematically lower than the other. In Table 5, last block, the estimate of $V(M_{\text{incr}\,1,2})$ is given for two increments; it differs imperceptibly from $V(\bar{l}_h - \bar{l}'_{h-1})$, as predicted.

Testing hypothesis B

In summary of this section, then, we may say there are two good ways of estimating the mean increment from one year to the next. The theoretically better way under most circumstances is by means of $M_h - {}_hM_{h-1}$, and this has a variance given by expression P. A simpler way, in most cases nearly as accurate if the percentage longitudinality is fairly high, is by dropping the cross-sectional cases completely and taking the estimate as the mean of the longitudinal element increment, as has been done in Dr. Thompson's paper. Under circumstances where the specified pattern of correlations from year to year does not hold, this estimate may indeed be theoretically preferable to the $M_h - {}_hM_{h-1}$ estimate. In circumstances where only an estimate of t_{h-1} to t_h *increment* is required, without an estimate of mean yearly value at t_h and t_{h-1}, the longitudinal-element-only estimate is the one of choice except in extraordinarily accurate work. since it is much faster to compute from the raw data than $M_h - {}_hM_{h-1}$. But if M_h and ${}_hM_{h-1}$ are themselves required there is little computational advantage on one side or the other.

The standard deviation of increment should be estimated from the actual t_{h-1} to t_h increments, i. e. from the longitudinal element alone. The deviation of a given child's increment from the estimated population mean is then compared to this estimate of standard deviation, and an opinion given as to whether or not the child comes from the

population the standards are based upon; if not, the child may well be abnormal.

Hypothesis C

There is thirdly hypothesis C, dealing with the child P's acceleration. Little need be said about this in principle, since the acceleration data bear the same relation to the velocity data as the velocity do to simple overall yearly figures. For any given acceleration from t_1 to t_3 one would have a longitudinal element of children present throughout the t_1 to t_2 period and throughout the t_2 to t_3 period, and a cross-sectional element of those present at t_1 and t_2 and absent at t_3, etc. From each of these elements one could get independent estimations. Note that this time there is yet a third element we might call distance-cross-sectional as opposed to velocity-cross-sectional. This third element consists of children present on only one of the three occasions t_1, t_2 and t_3. This element also can furnish an estimate of mean acceleration, though a poor one compared to the other two.

The weighted combination of these three estimates which gives the single best estimate has not been published though it could be worked out by the mathematical method Patterson uses for velocity estimates; no new principle is involved. In practice, however, considerable doubt remains as to how best to estimate the mean acceleration; at the moment acceleration is claiming little attention in the study of growth, but this will not always be the case. We could perhaps proceed as in velocity estimation and drop all the cross-sectional figures, both velocity-cross-sectional and distance-cross-sectional, and use the pure longitudinal series alone. However, it is probable that the estimate from velocity-cross-sectional data is not as greatly inferior to the longitudinal element estimate as it was in velocity data. It will be recalled that the relative importance of the estimate depends on the correlation coefficient for the longitudinal element, uniformly high in velocity data. The only incremental correlations (i. e. of increment $t_2 - t_1$ with increment $t_3 - t_2$) I can find are for the Harvard Growth Study data reported by Muhsam (7) and they are uniformly low; they range from zero to 0.6 for stature increments over the age span 7-15, and are practically uniformly zero over the same age span for weight.

Thus it seems that dropping the children not present on all three occasions will not lead to a very good estimate of acceleration.

The relative value of the hypotheses

The reporting of the data of a growth study relates to the hypothesis the study was designed to test. When an investigator embarks on a longitudinal study, he thereby declares that there is more value in hypotheses B and C than in A. For most purposes this seems a very reasonable belief. Growth, after all, is a form of motion, and motion is usually better characterised by velocity and acceleration than by end-results. The morphological events, too, only reflect biochemical processes, and if it is these that we wish to penetrate to, rate of growth seems a more appropriate characteristic to consider than size. Even from the pediatrician's point of view, it seems that one is more likely to interrupt abnormality early and successfully by studying velocities whenever serial examination makes it possible, rather than by measuring their end-result.

There is one further point that should be made explicit. We have been considering average yearly increments, and assuming these have some sort of validity as measures of velocity of growth. This they certainly have, but it must be remembered that when a measurement is rapidly decelerating during the year the average velocity over the whole year is not actually occurring at the mid-point of the year, but before it. The smaller the increment periods taken the better, as the nearer the calculated velocities then approach to the true curve, when the ripples due to seasonal differences have been ironed out. If the intervals are coarse, interpolations can be made by the methods described for example, in Yule and Kendall (8), or continuous curves can be fitted. There seems little point in fitting curves, however, especially to average increment figures, unless one has a definite hypothesis to be tested by such a procedure, and unless the sampling errors of the curves' constants can be calculated sufficiently well for the test of the hypothesis to be carried out. This point has been very clearly discussed by Sholl (9). Whether halfyearly periods should be used instead of yearly ones during the first few years depends partly on the data and partly perhaps, when the methods described above are used, on the pattern of the year-to-year and halfyear-to-halfyear correlations. If the latter fit the assumptions of the system better than yearly correlations, halfyears should perhaps be preferred.

Hypothesis D

Pure longitudinal series covering birth to maturity

So far we have considered only mean yearly values, mean year-to-year increments and mean accelerations. We have discussed how to make a test of whether a given child is of normal size at *a particular year,* or if he has grown normally *from one particular year to the next.* In testing these hypotheses we invoked only the necessity of having data for a given child which extended over two, or for acceleration, three years. It is true that our estimates of mean velocity will be slightly improved by having larger stretches of data on single children, but the improvement is really a trivial one. The question arises as to what advantage, if any, is gained by collecting longitudinal data that cover a span of over two years. Admittedly there are obvious and manifold reasons in the physiological, psychological and sociological fields for prolonging the observation of the particular child, but we are not immediately concerned with these. We want to know if we can derive additional information as to morphological changes from prolonged individual data, and if so, how such data should be handled so that we actually obtain this information.

Despite the number of longitudinal studies completed or in existence, relatively few serious attempts have been made to use their individual longitudinality. The difference of extended longitudinal studies over two-year ones is that a whole series of points are available for each child in the former, and these points can be characterised by some form of mathematical expression. A curve can be fitted to the points; and in the equation of this curve there will appear several parameters. The values of these parameters for each individual give the further information that we are seeking; the parameters give more than the simple velocity or acceleration at a particular age. Our previous hypothesis was that P's increment over a single year was normal; from 6 to 7 or from 7 to 8. We are substituting the more sensitive hypothesis D—that P's growth over a number of years, preferably from birth to maturity, is normal, as judged by the standards of his group. To test this hypothesis we need to fit curves to the series of points generated by single individuals and to do this usefully we need a good hypothesis about the type of curve, or group of curves, to be fitted, and a considerable number of points. Equations can be fitted to an individual's distance, velocity, or accelera-

tion curves, but the use of the distance curve seems preferable as its initial points are determined with only one measuring error, whereas velocity involves two. The fitted distance curves can be differentiated to give velocity and acceleration.

Suppose that the curve used is of a form which has three constants *a*, *b*, and *c*. For a given population there will be many such curves and thus a series of *a*'s, *b*'s, and *c*'s, whose means and variabilities can be calculated. The individual P's figures for *a*, *b*, and *c* can be then compared with the figures for the group. The procedure is arithmetically laborious, of course, but likely to produce very interesting results. Though innumerable curves of growth have been proposed, and considerable numbers of individual growth curves amassed, there does not seem to have been any attempt to carry through this whole approach on a practically useful scale. Merrell (26) in 1931 fitted individual weight curves for 46 rabbits, but only as an example in her mathematical treatment of the problem, which we will return to shortly. Jenss and Bayley (10) in 1937 emphasised the method and introduced an equation now being studied practically by Dr. Jean Deming (11). And in the course of a discussion on the growth of skull dimensions in early childhood, Count (13) has given a more complete example, in which a single individual's parameters are compared to the mean values for the group.

One of the things the method does is to eliminate time differences; that is, it allows for one child being ahead of another. Suppose, for example, we have two children who have adolescent spurts identical in form and intensity, but beginning at different times. If our curve is a well considered one, one or more of its parameters will characterise the slope of the spurt, another its peak intensity, and so on. These parameters will be equal for our two children, for they are unrelated to time. Two other figures, referring to the time at which the change in velocity began, will be different for the two children, and will characterise their chronological advancement or retardation. Such parameters are the appropriate way of investigating the exact relationship between the time of the beginning of the spurt and the peak velocity achieved, and similar questions. In fact, the outstanding use of the method is Shuttleworth's investigation on adolescence (1), dealing with this precise problem, amongst others.

Fig. 3 may help to make this clear. The two graphs are diagrammatic and drawn after the manner of Shuttleworth. The five solid curves in Fig. 3A are individual velocity curves plotted against time. In Fig. 3B

the same curves have been arranged so that their points of maximum velocity coincide. In each graph the dotted line represents the average curve of the five. It is obvious that this dotted curve in Fig. 3A is a very poor characteriser of average velocity; it smooths out the adolescent spurt, as pointed out some time ago by Boas (29) and Davenport (12).

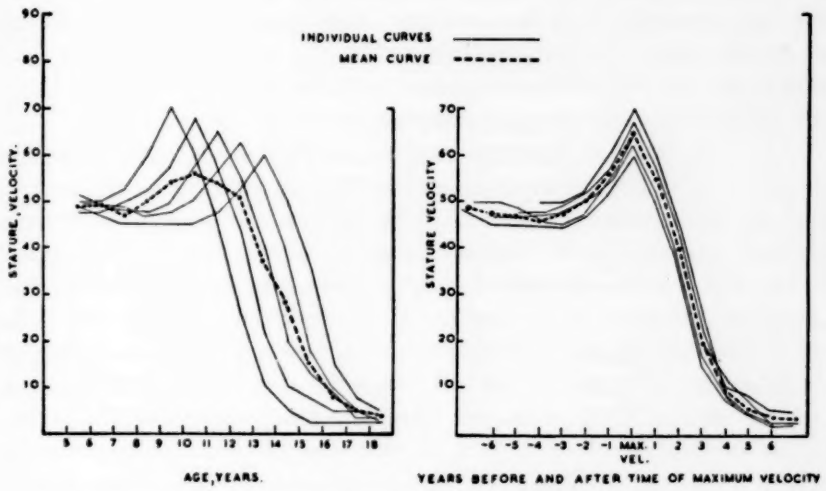

FIG. 3. THE RELATION BETWEEN INDIVIDUAL AND MEAN VELOCITIES DURING THE ADOLESCENT SPURT, WHEN THE INDIVIDUAL CURVES ARE CLASSIFIED BY AGE (FIG. 3A) AND BY THEIR ACCELERATION (FIG. 3B). AFTER SHUTTLEWORTH, 1937. Fig. 3A. Curves plotted against age; Fig. 3B. Same curves plotted with time of maximum velocity coinciding for each.

If individual P's spurt were compared to such an average, he would be thought quite abnormal. But in Fig. 3B, the average curve characterises the group quite nicely. This is Shuttleworth's graphical solution of the problem, a solution which proved exceedingly effective in his hands, but which is at best only semi-quantitative. The common practices of equating children for bone maturity age, or years before or after menarche, or growth of secondary sex characters are likewise attempts, more or less successful and only applicable over a short range of the growth period, to deal with this problem.

This brings us to the last point of this section. Besides fitting curves to an individual's points, one can fit a curve with the same type of equation, very often, to the mean values of a group, either of a cross-

sectional or of a longitudinal series. This *mean curve*, as we call it, will have certain values for a, b and c. But these mean-curve values will *not*, in general, be the same as the mean values of a, b and c obtained by fitting individual curves for persons X, Y and Z and averaging the a_X, a_Y and a_Z, the b_X, b_Y and b_Z, and the c_X, c_Y and c_Z. On the contrary, we can plot another curve which does have for its constants these average-of-individual values, and this curve is called the *mean-constant curve*, as opposed to the mean curve. In Fig. 3A the average curve is the mean curve of the five individual values at each age; in Fig. 3B the average curve approaches a mean-constant curve.

Merrell (26) has examined mathematically the relation between the mean curve and the mean-constant curve, to see how close a mean curve can get to the mean-constant curve, the one we really want. She found that it was possible to express the mean curve as the mean-constant curve plus a series of added terms. These added terms constituted an algebraic series, the first member of which involved the second order partial derivatives of the initial equation with regard to the constants a, b, c, etc. The second term of the series involved the third partial derivatives and so on. Thus the amount by which the mean curve differs from the mean-constant curve depends on the initial equation used, and when we have individual data available it can be evaluated. Merrell gives the examples of a parabola, where the mean curve and mean-constant curve are identical, and a logistic, where they differ considerably.

This result is of very considerable importance, and we will pause for a moment to consider further examples of its application. In Count's paper on skull growth already referred to (13), the average curve is given as a mean curve, not a mean-constant one; Count did not calculate each individual's a, b and c at all. He compares one particular individual's a, b and c however with the a, b and c of his mean curve. He implicitly assumes, then, that his mean curve is identical with his mean-constant curve. We can now judge whether he is justified in this assumption or not. The equation he fits to the points is $y = a + bt + c \log t$. The first partial derivatives of this with regard to first a, then b, then c, are 1, t, and $\log t$. The second are all zero. Consequently his mean and mean-constant curves *do* coincide. A little farther on in the same paper, he goes through exactly the same process using a dimension y plotted not against time t, but against another dimension x. This time his fitted equation is $y = ae^{bx}$, and this time the second partial derivatives are *not* zero. The mean curve he gives

is not identical with the mean-constant curve and the comparison of an individual's figures for a and b with those of the group curve involves a fallacy, the numerical importance of which is unknown.

Count's pioneering work in the mathematics of human growth curves provides us with another example, this time of a slightly different nature. Count has fitted a series of three curves to the cross-sectional birth-to-maturity data of stature (14). This mean curve has various characteristics, in particular three periods of acceleration,—prenatal, mid-growth and adolescent. Which characteristics of this curve characterise also the mean-constant curve, the one we are really interested in, and which came only from the sort of averaging process pointed out by Davenport? Merrell's work provides at least a partial answer. The first two of Count's curves are logarithmic, of the form $y = a + bt + c \log t$, and so no distortion from the mean-constant curve appears. From the beginning of adolescence onwards, however, the fitted curve is a logistic, the form Merrell examined closely. Her studies indicate that this mean curve of Count's will probably not produce any undulations that are not present in the curve of individuals, but it will misrepresent the degree of skewness of individual curves. Count's logistic for females is asymmetrical; Count's male mean curve is symmetrical and in this case the mean-constant curve is likely to be asymmetrical. The magnitude of this distorting effect of mean curves depends chiefly on the standard deviations of a, b and c, which remain unknown for this type of data.

PART TWO: TWO OR MORE MEASUREMENTS; PROPORTIONAL GROWTH

The second part of this paper is concerned with proportional or relative growth, and the relating of two or more measurements throughout the growing period. As a rule this has been done by plotting the ratio of two measurements as a curve in time; if the line rises, it is said that the upper member of the pair is growing faster than the lower member. And if the line obviously curls upwards, that is, has an increasing positive slope, it may be assumed that not only is the upper member growing absolutely faster than the lower but that it is growing faster and faster still, that it is accelerating away from the lower member. The assumption sounds reasonable, but it is an incorrect one. Ratios can rather easily involve elementary statistical fallacies where their user is unwary; some of these I have examined elsewhere (15). We

must consider now what is the appropriate way of reporting data for the various hypotheses of relative growth which one might wish to test.

Let us call the two dimensions l and m. As in the first part of this paper, there are three simple hypotheses we might be concerned with. *Hypothesis A* is that $l > m$ at a particular age; cross-sectional data for the age in question obviously answers this. The hypothesis might further state that $l > m$ by an amount which increases as time passes. On this hypothesis, then, $l/m > 1$ and increases as t increases, or $(l - m) > 0$ and increases as t increases. The relevant test is to plot l/m or $(l - m)$ against t. We then want to know if, say l_2/m_2 is greater than l_1/m_1, the subscripts referring to consecutive years. The procedure is the usual one of relating the differences between the two ratios to its standard error. Longitudinal data will be much more efficient than cross-sectional for testing the hypothesis because the standard error of this difference will then be much smaller for the reasons we have already discussed. We are comparing the velocities of the two dimensions and testing (albeit with bias, as discussed below) whether, from the period t_1 to t_2, l has a higher velocity than m. However, we are still essentially concerned with changes in form, not with changes in relative rate of growth.

Hypothesis B is that the *rates* of growth of the dimensions l and m are changing, one relative to the other. The hypothesis is that the velocity of l is greater than that of m, and is increasing relative to m's velocity as time passes. If it is true that biochemical processes are best searched for by considering rates of growth rather than size attained, then it is this hypothesis which has the most importance for us. If the fundamental control is of velocity, we must look for changes of velocity of l in relation to m; so long as the ratio of their velocity stays constant, we need not expect new chemical processes to be entering the picture, even if the external form *is* changing. The relevant function to plot is the ratio of the velocities, $dt/dt/dm/dt$, or using a dot above the letter to signify differentiation with regard to time, $\dot{l}/\dot{m}$. But I cannot find a graph of this function anywhere in the literature of human growth. Longitudinal data, of course, are needed, and the methods used for obtaining estimates of $\dot{l}_{12}$ and $\dot{m}_{12}$, $\dot{l}_{23}$ and $\dot{m}_{23}$ etc. are those described in Part 1 of this paper.

This notation makes it easy to see why we cannot in general arrive

at relative changes in velocity from looking at the slope of the l/m curve. The slope of the l/m curve is $\frac{d}{dt}\left(\frac{l}{m}\right)$ and

$$\frac{d}{dt}\left(\frac{l}{m}\right) = \frac{\dot{l}}{m} - \frac{l\dot{m}}{m^2} .$$

The slope of the l/m curve depends on the ratio of the two dimensions as well as on their relative velocities. This is illustrated by two examples in Fig. 4, where the solid lines represent $\dot{l}/\dot{m}$ and the dotted lines the slope of l/m, the raw data being the same in each instance.

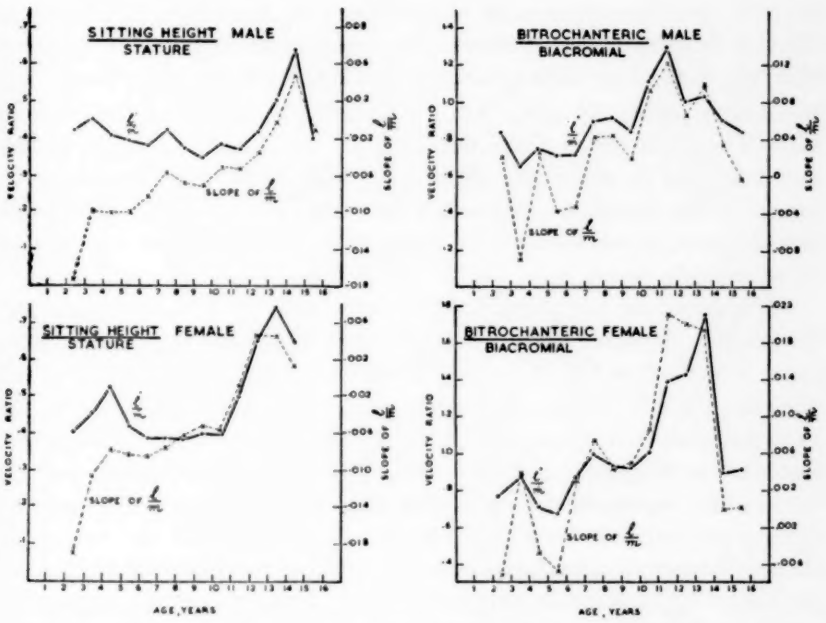

FIG. 4. COMPARISON OF CURVES OF $\dot{l}/\dot{m}$, THE RATIO OF TWO VELOCITIES, WITH THE SLOPE OF l/m. BRUSH FOUNDATION DATA, SIMMONS, 1944.

Solid line $\dot{l}/\dot{m}$, dotted line slope of l/m.

No vast discrepancy occurs in these two cases, but the slope of l/m fluctuates considerably more than $\dot{l}/\dot{m}$, and in some places the two plainly do not tell the same story. For example, in the male curve of sitting height/stature (upper left) the slope of l/m increases more or less uninterruptedly from 2 to 15 years. But the ratio of the velocities,

$\dot{l}/\dot{m}$, is constant, or even slightly decreasing, from age 2 till adolescence. On the other hand, $\frac{d}{dt}(l - m)$ does equal $(\dot{l} - \dot{m})$; if the absolute difference between the measurements is plotted, its slope does give the absolute difference of the velocities.

Hypothesis C is that l is accelerating more than m, and that the degree to which the accelerations differ increases with time. In this case $\ddot{l}/\ddot{m}$ or $(\ddot{l} - \ddot{m})$ must be plotted.

The function used to represent the relation between two dimensions

The function of l and m which would best represent the relation between the two dimensions seems never to have been discussed in relation to growth data. For a long time it has been tacitly assumed that the ratio l/m is a good one, but the assumption seems to me to be very insecurely grounded. Why not $(l - m)$, or $(l^2 - m^2)$, or $(l - \bar{l})^2 - (m - \bar{m})^2$, $(\log l - \log m)$ or $\log l/\log m$? Can we adduce any principle to guide our choice? The question applies equally, of course, to the relations between the velocities $\dot{l}$ and $\dot{m}$. (Allometry, discussed below, is an example of using a more complicated function of these, notably $\dot{l}/l/\dot{m}/m$.)

The situation has perhaps some slight analogy with two statistical procedures, the analysis of covariance and Studentization. The latter term is applied to the finding of a statistic relating, for example, two variables such that scale of measurement or absolute magnitude of the variables does not affect the statistic. The process of finding such statistics is of great importance in statistical estimation, and perhaps may be the criterion for our choice in the present situation. Presumably it was the desire to obtain a combined measurement independent of the scale which led to the adoption of the ratio l/m instead of the difference $(l - m)$. If l is a dimension with a mean of 50 cm. and m a dimension with a mean of 10 cm., it would appear at first glance that $(l - m)$ would be dominated by the value of l at the expense of m,—that $(l - m)$ would be more dependent on l than on m. But that this is not always so can easily be seen; suppose l were constant, then $(l - m)$ would depend more, and indeed solely, on m. Evidently the relative variabilities of l and m enter into the matter. Indeed, in general the correlation of $(l - m)$ and l is

$$r_{l-m,l} = (s_l - s_m r_{lm})/s_{l-m}$$

where s_l, s_m, and s_{l-m} are the standard deviations of l, m, and $(l - m)$, and r_{lm} is the correlation coefficient between l and m. The correlation of $(l - m)$ and m is similarly

$$r_{l-m,m} = (s_l r_{lm} - s_m)/s_{l-m}.$$

If $(l - m)$ is to be equally closely correlated with l and with m (naturally, with opposite sign) it is easy to see that we must have either

$$s_l = s_m$$

or

$$r_{lm} = 1.$$

A similar argument can be carried through for the ratio l/m. The correlation of l/m and l is approximately (see reference 15 for more detail)

$$r_{l/m,l} = \frac{V_l - V_m r_{lm}}{\sqrt{V_l^2 + V_m^2 - 2r_{lm}V_l V_m}},$$

and of l/m and m

$$r_{l/m,m} = \frac{V_l r_{lm} - V_m}{\sqrt{V_l^2 + V_m^2 - 2r_{lm}V_l V_m}},$$

where V_l, V_m are the coefficients of variation of l and m. For equally close correlation of l and m with l/m we must have

$$V_l = V_m,$$

that is, equal coefficients of variation. In actual anthropometric work we rarely have either equal standard deviations or equal coefficients of variation. It is true that in some groups of anthropometric measurements the dimensions with bigger means have bigger standard deviations (for example, stature, arm length, biacromial diameter) but increase in standard deviation does not anywhere near keep pace with increasing mean. Consequently the measurements with the largest means, such as stature, have the lowest coefficients of variation. Taking l/m instead of $(l - m)$ to represent the l, m relationship thus grossly overcorrects for the variances of the two measurements not being equal. We are out of the frying-pan into the freezer; somewhere in between lies the temperate zone. For instance, since stature has the lowest coefficient of variation of all the usually taken adult measurements, none of the per-stature indices (such as biacromial/stature), so dear to the tradi-

tional anthropologist, is independent of its two components. If the coefficient of variation of the other measurement is not equal to the correlation coefficient between the two, the ratio will increase or decrease as stature increases and decreases. For most per-stature ratios in fact, the ratio of the coefficients of variation is less than the correlation coefficient, and the per-stature index decreases as stature increases. This effect, it seems to me, nullifies in many or most circumstances, any usefulness the index may be supposed to possess.

One possible way out of this difficulty is to relate two dimensions always in terms of their *standard scores*, such scores being of the form $(l - \bar{l})/s_l$, the deviation from the mean divided by the standard deviation. If the distribution of the measurements concerned is normal, this Studentizes them, i. e. makes them independent of scale. The procedure has been much used by psychologists and, in a slightly modified form, by the distinguished Italian school of constitutional anthropology. Such scores are often referred to as z-scores, and z_l and z_m may be related to each other either by subtraction or division without the result being dependent on l or m, since the variances and the means of all z-scores are equal.

Translated into velocity terms, this approach would mean reporting increments of two dimensions in standard scores, that is as deviations from the mean increment divided by the standard deviation of increment. The principle could be extended, presumably, beyond two dimensions and might indeed provide the best means for comparing and plotting the velocities of several dimensions at once. So far as I am aware, this has never been done for growth data. Unfortunately the lack of any published standard deviations of increment except for height and weight (whose distribution is too skewed to serve) prevents an example being presented here.

A second statistical analogy is with the analysis of covariance, though the problem for which the analysis of covariance might be appropriate is different from the last. We might frame the question of l, m velocity relationship as follows: did l increase from age 8 to 9 more than m increased in the same year, each relative to their increments during the preceding 7 to 8 year? Proceeding without biometrical guide, we might try to answer this by comparing either the difference of l increments $(l_{8,9} - l_{7,8})$ with the difference of m increments $(m_{8,9} - m_{7,8})$. Or more probably by comparing the ratio of l increments l_{89}/l_{78} to the ratio of increments m_{89}/m_{78}. Neither procedure would give the proper answer;

we have no knowledge that either allows correctly for the 7 to 8 year increments. The analysis of covariance specifically deals with this problem; it works out the regression of the l_{89} values on the l_{78} ones and corrects the l_{89} mean for this and likewise corrects the m_{89} mean for the m_{78}, m_{89} regression. The two corrected means are then compared. The procedure can be generalised to cover more situations likely to arise in handling growth data. Corrections can be made for more than one previous value. Corrections could be made for another dimension n, instead of, or in addition to, correcting for previous values. Comparisons could be made between more than two measurements at a time. Thus if one suspected that in a certain year a particular dimension was growing faster than a group of others (say leg length relative to body widths from 8 to 9) an analysis of covariance could be done, allowing for either previous increments, a third group of bodily dimensions or both.

More than two dimensions

Ratios such as l/m compare only two dimensions at a time, and though it is true that a large number of such comparisons may be made to yield, subject to the bias mentioned above, a fairly general picture of the whole body, a technique for comparing simultaneously more than two dimensions would be exceedingly welcome. The standard score technique should accomplish this, and perhaps the covariance technique to some extent in the questions it is designed to answer. Another, and very interesting, approach would seem to be through the intercorrelation of the increments. If the velocity curves of two dimensions were exactly parallel, their correlation would then be 1, and deviations from parallelism, at least under certain circumstances, will decrease the correlation. A matrix of such correlations for a large number of measurements can be constructed and factor-analysed. Studies still in progress using mean increments indicate that this is quite a fruitful method. There are various difficulties to be faced, chief amongst them the difficulty or impossibility of estimating the standard errors of the correlation coefficients, since the velocity curves are autoregressive. There is also the objection that a matrix of mean intercorrelations is not identical to a matrix of the mean values of individual correlation matrices. The former is relatively easy to compute, the latter almost prohibitive; the relations between the two remain to be worked out.

Another approach is the method of co-ordinate transformations due to D'Arcy Thompson (25). There would have to be, I imagine, several

successive transformations to cover the growth of a human from birth to maturity. There might be a first general one stretching from birth onwards and altering the infants proportions to those of the adult. The transformations representing the shape factors would interfere with this general transformation, more at some periods such as the mid-growth spurt and adolescence, than at others. The final transforming co-ordinates would be rather wavy, suggesting the analogy of a compound energy wave before Fourier analysis. One can imagine a system of electronic lenses which would produce from a baby's dimensions, or even, shall we say, from a 3-month-old foetus, pictures of adults of every conceivable build according to how the lenses were set. (One can imagine too the chief use to which such machines would undoubtedly be put.)

Multiplicative velocity

We have dealt so far with simple velocities $\dot{l}$ and $\dot{m}$. It has often been proposed, however, that rate of growth may be more usefully described by using velocities proportional to the size of the dimension that is growing. Such a velocity is $\dot{l}/l$, and has usually been called the *specific growth rate*. This seems to me an unfortunate name since it becomes very confusing in circumstances where different zoological species are being considered; the word "specific" then gets used in two quite different senses at the same time. I propose to call this formulation *multiplicative velocity*, since the idea of its use is rooted in the conception that a cell divides into two parts, each of which can divide into two further parts, and so on. Thus growth might be expected to be geometrically progressive, with $\dot{l}$ proportional to l. If so, it is argued, it would be better to divide $\dot{l}$ by l when comparing it with the growth of dimension m. Otherwise, if l is bigger than m, and its constituent cells are dividing at the same rate, l will appear to be growing much faster since it has progressively more and more cells to divide. But if $\dot{l}/l$ is compared with $\dot{m}/m$ then the actual rates of division of the cells of each are compared, and this is what is held to be important. However only in the simplest biological situations does each of the products of all divisions actually go on forthwith to division itself; differentiation of cells occurs and mitosis and the differentiated functions are thought to be in general somewhat opposed one to another. In a bacterial culture derived from a single cell the daughter cells all divide again at approximately the same rate, say once every 20 minutes. This system is a homogeneous one. But in a multicellular organism there are many

different sorts of cells, each dividing, very likely, at different rates and providing a thoroughly heterogeneous system. Even the heterogeneity is not constant, since different sorts of cells very probably differ in reproduction rate from one period of growth to another (30). Only during foetal life and during the general growth spurt at adolescence does a master chemical reaction appear which at any rate to some extent causes a general cellular reproduction. For similar reasons, metazoan growth cannot be simply compared to the growth of a human population, which is again unimodal and moderately homogeneous in reproduction rate.

It must be realised of course that there is nothing necessarily wrong or incorrect in using either simple velocity or multiplicative velocity curves; the choice depends on the relative usefulness in the case in hand, and it is this that is debatable. At least in the growth of external dimensions in the human, $\dot{l}$ is by no means directly proportional to l. If it were, the correlation between the t_1 to t_2 increment and the t_1 size would be unity or close to it, so long as the t_1 to t_2 interval were not too large. Data on this point are available (6, 16, 17, 27) for height, weight and hip width, and in the case of the first the correlation is even below the level for significance, from birth to the beginning of the adolescent spurt. During adolescence some correlation exists, first positive, then negative. Weight increments do correlate to some extent with weight achieved, but quite apart from the association not being tremendously high, weight is a thoroughly atypical measure of growth in external dimensions or form. This is scarcely surprising since weight is not an external bodily dimension; its inclusion on the same footing as the linear measurements in correlation tables and so forth seems to me most questionable.

If multiplicative velocity is adopted for plotting, then the logarithms of the raw figures should be tabulated, and the t_1 to t_2 increments in logs plotted, since $\dot{l}/l = d(\log l)/dt$. Note that this is not the same thing as plotting the log of the straight increment, or plotting increments on log paper. Turning the original observations into logs is a more legitimate way of calculating the multiplicative velocity than taking the t_1 to t_2 increment and dividing by either the t_1 value, the t_2 value or the average of the two values, since the t_1 to t_2 velocity may not be a straight line (see 18, 19).

In Fig. 5 are plotted straight velocity and multiplicative velocity curves for some of the Brush Foundation measurements (5). These mean increments are subject to the errors of cross-sectionally reported

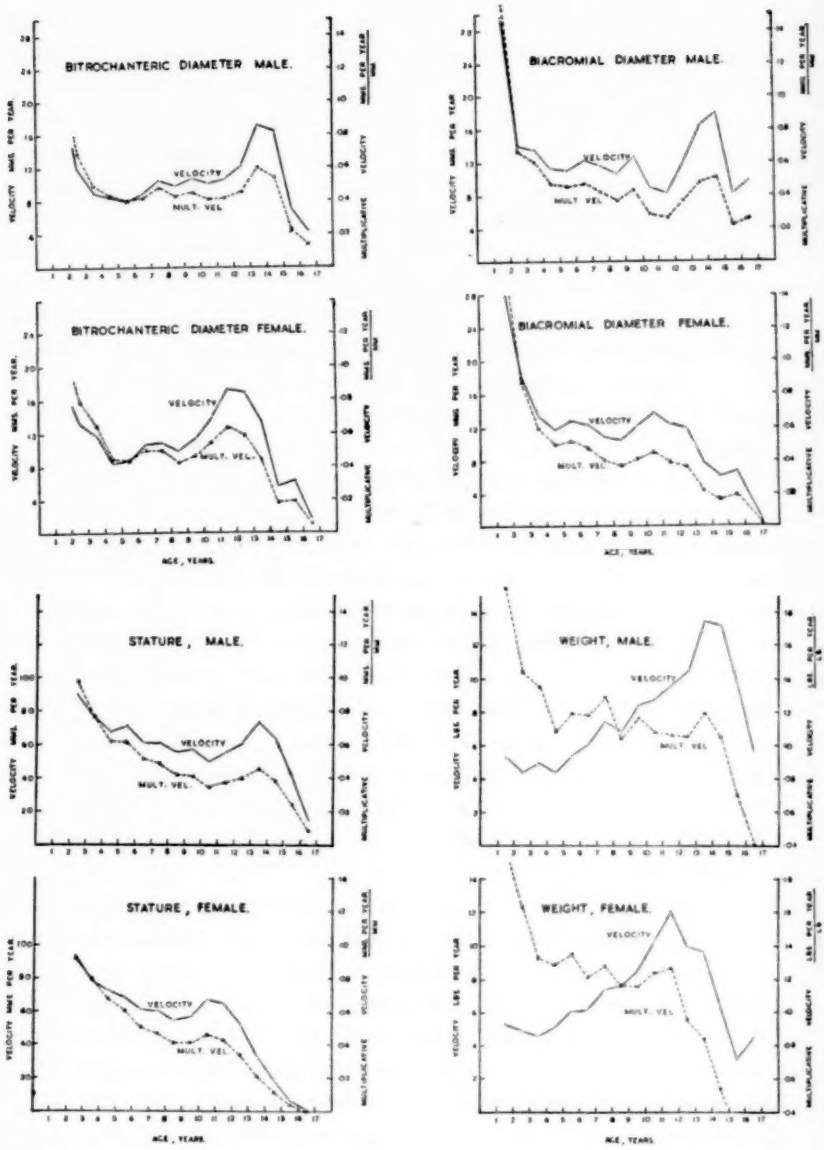

FIG. 5. COMPARISON OF VELOCITY AND MULTIPLICATIVE VELOCITY CURVES. BRUSH FOUNDATION DATA, SIMMONS, 1944.

Solid line velocity, dotted line multiplicative velocity.

longitudinal series discussed above, but this does not affect the present general argument. It is at once apparent that in the graphs of biacromial diameter, bitrochanteric diameter, and stature, the two lines are practically parallel. The effect of using multiplicative velocity instead of simple velocity is to decrease the magnitudes of the mid-growth and adolescent spurts slightly, a result that follows from these periods being the time of highest correlation of l and $\dot{l}$. In the first year or two the reverse happens, and the multiplicative curves emphasise the deceleration more than do the others. The differences between the two however are really quite trivial. In the case of body weight, the last of the four graphs, the difference is of an altogether different order. Here the multiplicative curves have the same effect as before, but on a much larger scale. The simple velocity curve of weight looks very different from the curves of the linear dimensions, but the multiplicative curve belongs very much more with the linear family. Perhaps there really is some advantage in using the more complicated form of velocity for weight; there seems none for the linear dimensions. The figure emphasises how rash it is to generalise about growth, and more particularly about growth of form, from a study of the growth in weight of an organism or an organ.

There seems therefore no very good reason for making the transformation, as it is called, from ordinary measurements to logarithms in work upon external human dimensions. It merely adds a certain amount of complication to the subject, and diminishes relatively just those phases of growth that claim the greatest attention. In statistical work there are several good reasons for transforming given data into logs, chief amongst them skewness with a long tail to the right which gets corrected towards normal by the transformation,—a situation arising when the magnitude of experimental errors increases as the size of the object measured increases, and when, in repeated samples, the standard deviations of the samples are correlated with their means. These things do not happen to any extent in human external dimension data, though they may with measurements of such things as subcutaneous tissue thickness.

Allometry

If multiplicative velocity is accepted, however, as preferable in a particular instance, then the relative rates of growth of dimensions l and m are given by $\dot{l}/l/\dot{m}/m$ or

$$\frac{d(\log l)/dt}{d(\log m)/dt}.$$

This time it is this ratio that is plotted and examined as to whether it increases, decreases or remains constant. This ratio is equal to k in the allometry formula $l = bm^k$, developed by Huxley (20) and others. In fact the allometric formulation of the relative growth of two parts of an organism was specifically designed in accordance with this postulate. If we take logs of the formula and then differentiate with respect to t, we have

$$\log l = \log b + k \log m$$

$$\frac{d}{dt} \log l = k \frac{d}{dt} \log m$$

$$\frac{d(\log l)/dt}{d(\log m)/dt} = k.$$

It was originally contended that k (also called by some authors α) remained constant in many or most circumstances, or at any rate changed only once or twice, and that abruptly, during the period of growth. It is now clear that this can no longer be maintained; in most mammalian situations k varies with the passage of time. A graph of k against time simply plots the ratio of two multiplicative velocities: such graphs are to be found, for example, in the exceedingly valuable article by Richards and Kavanagh, which contains the best discussion to date of the whole problem of allometry and relative growth (21).

In Fig. 6 several graphs of this sort are presented for human dimensions, using the Brush Foundation data, subject to the inaccuracies already discussed. The solid lines plot simple velocity ratios $\dot{l}/\dot{m}$ which I have called j, and the dotted lines multiplicative velocity ratios, or Huxley's k. That the two formulations tell very much the same story is again obvious. Equally so is the fact that k varies with time and has its own particular growth curve according to the dimensions considered. The upper four graphs for instance, show that at adolescence the bitrochanteric diameter grows more rapidly, both in the simple and in the multiplicative sense, than does stature or biacromial diameter. There is a hint also that bitrochanteric diameter may grow more rapidly than biacromial during the mid-growth spurt. The stature/weight lines (in the lower right hand quadrant) decrease fairly regularly. Only in the sitting height/stature, in these examples, does k seem to remain sensibly constant for a few years, at least from 5 to the onset of adolescence.

The last example gives us some insight into the relations between allometry and the results of factor analysis of the adult dimensions.

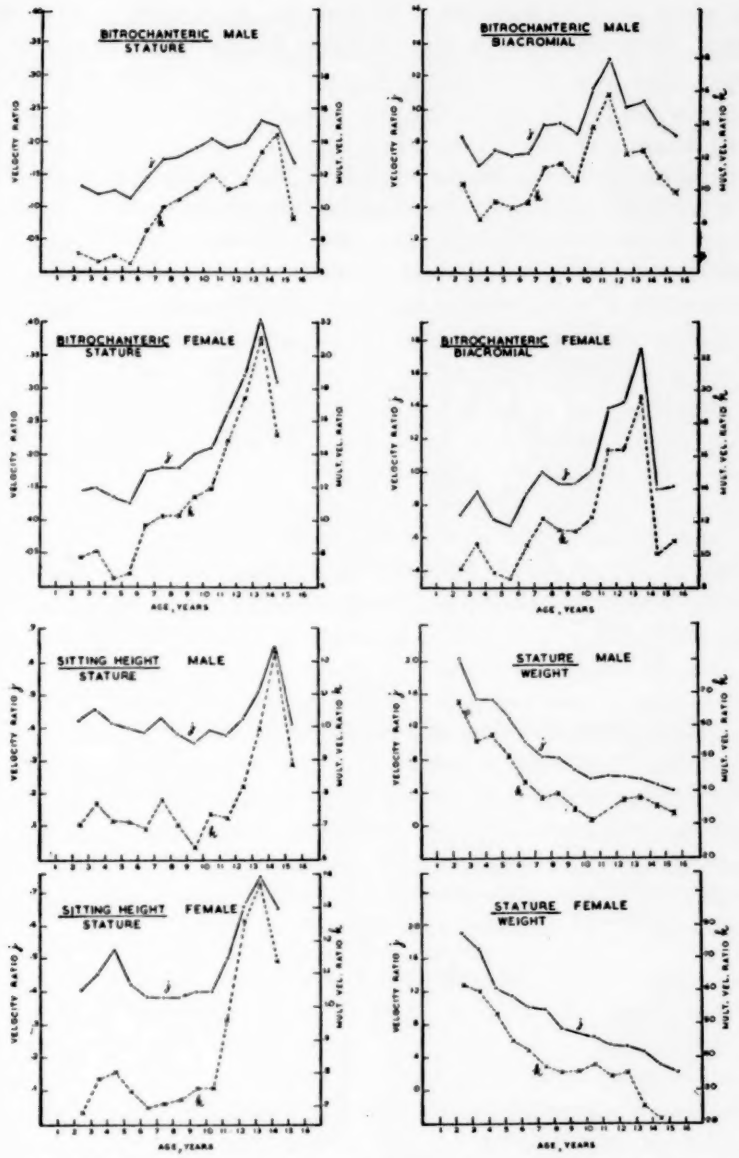

FIG. 6. CHANGES WITH AGE IN VELOCITY RATIOS *j*, AND IN MULTIPLICATIVE VELOCITY RATIOS, THE *k* OF THE ALLOMETRY EQUATION. BRUSH FOUNDATION DATA, SIMMONS, 1944.

Solid line velocity ratio, dotted line multiplicative velocity ratio.

The factor analysis leads to a general size factor and several shape factors. It has been shown that the shape factors reflect periods during growth when one set of bodily dimensions grows more rapidly than other sets (22). At adolescence, for example, there is a large spurt of sitting height and little of leg length: and at just this time a particular group factor for sitting height emerges in analyses done for each yearly age group (23). The graph of sitting height/stature shows the enormous increase in the value of k just when this factor appears. So long as no shape factors are operative, it seems that k stays constant, at least for certain pairs of more or less comparable dimensions. In fact, k was designed as a constant specifically on the assumption that changes in the rate of self-multiplication affect all parts of the body equally (24). It seems, then, that so far as k is constant, it measures the general factor of size, in the multiplicative system; and its departures from constancy measure the amount of differential development of the various shape factors, from between the two dimensions considered. Thus k stays more nearly constant for arm length/leg length where the advent of the leptosomic factor affects both similarly, than for arm length/chest breadth, where, though the general factor operates on both, the leptosomic does not.

SUMMARY

1. This paper discusses the biometrical handling of growth data, particularly those from a mixed longitudinal series, i. e. a series of children some but not all of whom leave the study each year, and some of whom, of various ages, arrive newly each year. Part 1 discusses the case of the single dimension such as stature; Part 2, the relations between two and more dimensions. An appendix gives detailed numerical calculations of the statistics recommended in Part 1.

2. *Part 1.* Data are collected to test hypotheses. Three possible hypotheses are: (A) child P is abnormal in stature at a given age; (B) child P grew abnormally fast from one given year to the next, i. e. had an abnormal velocity of stature; (C) child P accelerated abnormally in stature. The general method whereby the physician or biologist tests each of these hypotheses is outlined, P's deviation from the mean of the population of which he is thought to be a member being compared with the standard deviation of the population. It is emphasised that the data from normal children on which our standards are based are essentially sample data from larger populations.

3. The way to estimate the population mean for testing hypothesis A is detailed. Mixed longitudinal data must not be treated the same as cross-sectional data, or a large decrease in efficiency of estimation results. The various formulae for estimating population mean due to Patterson are set forth and their use discussed.

4. The estimation of mean yearly increment needed for testing hypothesis B is discussed. The usual practice of reporting growth study data only in terms of overall means and standard deviations at each year is strongly condemned, and it is shown how a sensitive test of hypothesis B cannot be made using standards reported in this way.

5. The formula for the best estimate of population mean increment is discussed, and it is shown that the estimate made by taking only the twice-measured longitudinal element of a mixed series and dropping the added cross-sectional cases entirely is nearly as efficient as the best estimate under the usual conditions of human growth studies, and simpler to compute.

6. The advantage gained by having a series of measurements on a single child extending over a considerable number of years is discussed, and the conclusion reached that it can only be realised by fitting a curve to the individual's measurements. The parameters of this curve are then compared with similar parameters for other children, thus testing a fourth and more critical hypothesis—that the child P grew normally over a long span of years. The point is made that the curve fitted to the mean measurements of a group of children will not in general have for its parameters the mean values of the parameters of the individuals' curves.

7. If the biometrical procedures outlined are used in reporting growth data, considerably more information will become available from growth studies than has hitherto appeared. However, because no biometrical procedure is final and the accumulation of longitudinal data is exceedingly costly and time-consuming, *every effort should be made to publish or put on easily available record the original raw data* of such studies, however excellent the biometrical workup.

8. *Part 2.* The reporting of the relation between two dimensions as the ratio of one to another is criticized. The analogy of Studentization is used in discussing what function of the two dimensions best charac-

terises their relation, and it is suggested that this function is one that is independent of the scales of measurement of either dimension. Standard or z-scores are such a function and their use is recommended. Analysis of covariance technique might also be used in certain situations.

9. The relations between increments of more than two dimensions considered simultaneously is discussed; factor analysis of increment inter-correlations, and a successive series of co-ordinate transformations are mentioned as possible approaches.

10. The reporting of increments as multiplicative velocity curves $(1/l)(dl/dt)$ is criticised and graphs given from human data showing the parallelism of such curves to ordinary velocity curves. The point is made that in general correlations between yearly increment and size at the beginning of the year are low; and the biological basis of multiplicative reporting in the human is called into question.

11. The question of allometry is discussed, and graphs given showing that in the human k, the so-called constant of the allometry equation, and the ratio of two multiplicative velocities, is far from being constant during growth and plots a regular curve which closely parallels the curve for the ratio of ordinary velocities of the two dimensions considered. The relation of allometry in growth and the factor analysis of adult physique is briefly discussed.

12. In the appendix a sample lay-out for reporting mixed longitudinal data is given, with detailed numerical calculations of the various estimates mentioned in Part 1.

ACKNOWLEDGEMENTS

Besides my debt to M. J. R. Healy and H. L. Patterson mentioned already, I wish to acknowledge the helpful suggestions of A. R. Jonckeere, and Professor Winsor's advice as to notation. I wish also to thank Professor Henry Barcroft for his interest and encouragement, and for the provision of every possible facility. I am also most grateful for the skilled computational assisatnce of Mr. R. H. Whitehouse, and particularly indebted to the friendly advice of Dr. C. E. Palmer.

REFERENCES

1. F. K. Shuttleworth. *Monographs Soc. Res. Child Develop.*, 2, No. 5, 1937. Sexual maturation and the physical growth of girls age six to nineteen.
2. E. B. Wilson. *Science*, 95: 112, 1942. Norms of growth.
3. H. Thompson. *Human Biol.*, 23: 75, 1951. Data on the growth of children during the first year after birth.
4. C. E. Palmer. *U. S. Pub. Health Rep.*, 48: 211, 1933. Seasonal variation of average growth in weight of elementary school children.
5. K. Simmons. *Monographs Soc. Res. Child Develop.*, 9, No. 1, 1944. The Brush Foundation study of child growth and development. 2. Physical growth and development.
6. K. Simmons and T. W. Todd. *Growth*, 2: 93, 1938. Growth of well children: analysis of stature and weight, 3 months to 13 years.
7. H. V. Muhsam. *Human Biol.*, 19: 260, 1947. Correlation in growth.
8. G. U. Yule and M. G. Kendall. An Introduction to the Theory of Statistics. 13th ed. (London), Griffin, 1946.
9. D. Sholl. *Proc. Roy. Soc. B.*, 135: 243, 1948. The quantitative investigation of the vertebrate brain and the applicability of allometric formulae to its study.
10. R. M. Jenss and N. Bayley. *Human Biol.*, 9: 556, 1937. A mathematical method for studying the growth of a child.
11. J. Deming. Personal communication.
12. C. B. Davenport. *Cold Spring Harbor Symposia*, 2: 203, 1934.
13. E. W. Count. *Human Biol.*, 14: 143, 1942. A quantitative analysis of growth in certain human skull dimensions.
14. E. W. Count. *Human Biol.*, 15: 1, 1943. Growth patterns of the human physique: an approach to kinetic anthropometry.
15. J. M. Tanner. *J. Appl. Physiol.*, 2: 1, 1949. Fallacy of per-weight and per-surface area standards and their relation to spurious correlation.
16. C. E. Palmer and L. J. Reed. *Human Biol.*, 7: 319, 1935. Anthropometric studies of individual growth. 1. Age, height and growth in height, elementary school children.
17. C. E. Palmer, R. Kawakami and L. Reed. *Child Develop.*, 8: 47, 1937. Anthropometric studies of individual growth. 11. Age, weight and rate of growth in weight, elementary school children.
18. R. A. Fisher. *Annals Appl. Biol.*, 7: 367, 1920. Some remarks on the methods formulated in a recent article on "The quantitative analysis of plant growth."
19. H. Gray. *Human Biol.*, 13: 306, 1941. Individual growth-rates: from birth to maturity for fifteen physical traits.
20. J. Huxley. Problems of Relative Growth. (*London*) 1932.

21. O. W. Richards and A. J. Kavanagh. The analysis of growing form. *In*: Essays on Growth and Form, presented to D'Arcy Wentworth Thompson, ed. by W. E. Le Gros Clark and P. B. Medawar. (*Oxford*) 1945.

22. J. M. Tanner. *Proc. Roy. Soc. Med.*, 40: 310, 1947. The morphological level of personality. Reprinted with additional figures in *Yearbook of Physical Anthropology* 1947 (*Ann Arbor, Michigan*) Edwards.

23. F. A. Mullen. *Child Develop.*, 11: 27, 1940. Factors in the growth of girls.

24. E. C. Reeve and J. S. Huxley. Some problems in the study of allometric growth. *In*: Essays on Growth and Form, presented to D'Arcy Wentworth Thompson, ed. by W. E. Le Gros Clark and P. B. Medawar. (*Oxford*) 1945.

25. D'Arcy W. Thompson. On Growth and Form. 2nd ed. (*Cambridge*) 1942.

26. M. Merrell. *Human Biol.*, 3: 37, 1931. The relationship of individual growth to average growth.

27. H. V. Meredith and L. J. Carl. *Child Develop.*, 17: 157, 1946. Individual growth in hip width: a study covering the age period from 5 to 9 years based upon seriatim data for 55 non-pathologic white children.

28. F. Boas. *Human Biol.*, 4: 307, 1932. Studies in growth.

29. F. Boas. *Science*, 19: 256, 281; 20: 351, 1892. The growth of children.

30. J. Gray. *Brit. J. Exp. Biol.*, 6: 243, 1929. The kinetics of growth.

31. F. Yates. Sampling methods for censuses and surveys. (*London*) Griffin, 1949.

32. J. W. Tukey. *Human Biol.*, 20: 205, 1948. Some elementary problems of importance to small sample practice.

33. H. V. Meredith. *Univ. of Iowa Studies in Child Welfare*, 11, No. 3, 1935. The Rhythm of Physical Growth.

34. H. D. Patterson. *J. Roy. Stat. Soc. B.*, 12: 241, 1950. Sampling on successive occasions with partial replacement of units.

35. H. V. Meredith. *Growth*, 11: 1, 1947. Growth of upper extremities of *homo sapiens* from birth through adolescence.

36. M. M. Maresh. *Amer. J. Dis. Child.*, 66: 227, 1943. Growth of major long bones in healthy children.

37. E. B. Wilson. *Proc. Nat. Acad. Sci.*, 21: 633, 1935. Heights and weights of 275 public school girls for consecutive ages 7 to 16 years inclusive.

APPENDIX

NUMERICAL EXAMPLES

In this appendix two examples are given of the full work-up, as detailed in Part 1, of data on a single measurement. The figures are presented in sufficient detail for the computations to be followed by investigators with a minimum of biometrical knowledge, or by secretary computers.

The first example is taken from Meredith's data (35) on the growth of the upper limb (acromion process to tip of middle finger, arm extended down the side of the body) in girls from age 3 to 12. The subjects were healthy children attending the University of Iowa Preschool Laboratories and Elementary School during the years 1937-1946; they were mostly from the professional and managerial socio-economic group. These data are rather highly longitudinal and provide a somewhat simplified application of the methods of Part 1 in that, as can be seen from Table 4, new children are only joining during the first three years. From 6 to 8 the data are completely longitudinal, and at no year do new children join *and* old children leave.

The second example is drawn from Maresh's study of the growth in length of the humerus (36). The bone was measured by X-rays, the focus-film distance being $7\frac{1}{2}$ feet throughout, and the humerus measured from one epiphyseal line to the other along the long axis of the shaft. The children were healthy girls attending the Child Research Council's study of growth at Denver, Colorado. The data are for the most part reported over 6-monthly intervals and I have followed this in the analyses, omitting in Table 6 a few values which related to intervening 3-monthly measurements. As can be seen from the table, this data is much more fragmentary in the longitudinal sense than that of Meredith, and provides a more complete, and slightly more complicated application of the methods of Part 1.

The publication of actual raw data in growth studies is still a rare event, and one must be most grateful to the authors for making these series generally available and even more so for the foresight of the editors who managed to get them published.

UNIVERSITY OF MICHIGAN LIBRARIES

Notation

We repeat here, with some additions and omissions, the notation given earlier.

For all subjects present at t_h:

N_h = number of subjects;

Σx_h, Σx_h^2 = sum and sum of squares of measurements;

m_h = mean of all measurements;

$v_h = \Sigma(x_h - m_h)^2/(N_h - 1)$ = variance of all measurements;

$s_h = \sqrt{v_h}$ = standard deviation.

For subjects present at t_h who were also present at t_{h-1}:

L_h = number;

Σl_h, Σl_h^2 = sum and sum of squares of measurements;

$\bar{l}_h$ = mean; [5]

$V(l_h)$ = variance.

For subjects present at t_h but not present at t_{h-1}:

$$C_h,\ \Sigma c_h,\ \Sigma c_h^2,\ \bar{c}_h,^5\ V(c_h).$$

We shall also need to consider the measurements of subjects at t_{h-1} who were also present at t_h. We denote these by primed symbols

$$L'_{h-1},\ \Sigma l'_{h-1},\ \text{etc.}$$

Finally, we shall need the covariance of the measurements of subjects present at both t_{h-1} and t_h:

$$\text{cov}(l'_{h-1}, l_h) = \Sigma(l'_{h-1} - \bar{l}'_{h-1})(l_h - \bar{l}_h)/(L_h - 1).$$

In actual computation, we work from the relation

$$\Sigma(l'_{h-1} - \bar{l}'_{h-1})(l_h - \bar{l}_h) = \Sigma l'_{h-1} l_h - \Sigma l'_{h-1} \Sigma l_h / L_h .$$

Example 1. The layout of data and basic calculations in Tables 4 and 5 is the one we have found most convenient in computation. The arrangement of the raw data follows Meredith directly and requires no comment. The next two blocks, sums and sums of squares of values, are best taken together with a calculating machine. Such a figure as 831.4 for Σl_h, for instance, represents the sum of the age 5 values of those children who were present at both ages 4 and 5. Similarly, the

[5] In Tables 5 and 7, l_h is used in place of $\bar{l}_h$; c_h in place of $\bar{c}_h$.

TABLE 4

TABLE 4

Length of upper limb (cms.) age 3 to 12 years in healthy girls; data from Meredith 1947

SUBJECT NUMBER	AGE IN YEARS 3	4	5	6	7	8	9	10	11	12
895	40.5	43.7	46.9	50.2	53.1	55.2	..	..	..	..
2450	41.0	44.7	48.0	51.1	54.0	56.6	..	..	..	..
4254	40.2	44.4	48.0	51.2	54.1	57.0	..	..	..	..
4587	41.4	44.7	48.3	51.7	55.0	57.6	..	..	..	..
4746	38.9	42.4	45.5	48.7	51.4	54.0	..	..	..	..
4865	38.6	42.0	45.4	48.7	51.7	54.2	..	..	..	..
8192	39.3	42.8	46.2	49.5	52.4	55.3	..	..	..	..
3320	40.8	43.6	46.6	49.7	52.7	55.6	58.4	..	..	..
9734	38.4	41.1	44.2	47.3	50.2	52.6	54.4	..	..	..
1013	38.4	42.1	45.2	47.5	50.0	52.4	54.6	56.3	..	..
1157	40.0	43.4	47.0	49.7	51.9	55.2	58.2	60.7	..	..
2333	38.9	41.0	43.8	46.9	50.0	52.9	55.8	58.2	..	..
4937	39.5	43.4	46.9	50.2	53.4	56.1	58.9	61.8	..	..
1169	..	40.7	44.7	47.9	50.5	52.7	54.4	..	..	..
424	..	40.3	43.1	45.9	48.7	51.3	53.8	56.7	..	..
1673	..	43.9	47.1	50.2	53.3	56.3	59.1	61.2	..	..
4224	..	46.4	49.6	52.6	56.1	59.4	62.9	65.3	..	..
9356	..	41.4	44.9	48.5	51.8	54.6	57.3	60.0	62.7	..
4879	..	..	46.8	50.0	53.1	55.9	58.7	61.1	..	..
383	..	..	43.5	46.7	49.9	52.9	55.5	57.9	60.3	..
2049	..	..	46.6	49.8	52.8	55.7	58.8	62.1	66.4	..
2178	..	..	47.1	50.6	53.7	56.8	59.7	62.2	65.4	..
2678	..	..	46.3	48.7	51.2	53.9	56.8	60.1	63.0	..
7535	..	..	45.7	48.7	51.3	53.5	55.7	58.2	60.6	..
2318	..	..	46.5	49.5	52.7	55.5	58.1	60.5	63.0	66.6
4749	..	..	44.9	48.0	51.0	53.7	56.4	59.4	62.7	67.4
5413	..	..	45.9	49.2	52.5	55.3	57.8	60.6	64.0	67.8
6207	..	..	46.4	48.6	51.0	53.8	56.6	58.9	60.9	62.6
753	..	..	..	47.1	49.2	51.7	54.8	58.0	60.7	..
4158	..	..	..	46.0	48.9	51.4	53.7	56.8	61.1	..
4398	..	..	..	46.7	49.5	52.3	54.9	57.3	59.6	..
211	..	..	..	48.0	50.6	53.2	55.8	58.4	61.3	65.3
393	..	..	..	47.2	50.2	52.9	55.5	57.6	60.4	64.0
472	..	..	..	52.2	55.3	58.1	61.1	63.8	67.2	70.5
2147	..	..	..	49.5	52.7	55.6	57.9	60.0	62.1	65.1
7732	..	..	..	48.1	51.0	53.7	55.9	57.9	59.8	62.5
$\Sigma\, l_h$	0	559.3	831.4	1377.3	1866.9	1964.9	1651.5	1551.0	1121.2	591.8
$\Sigma\, c_h$	515.9	212.7	459.7	384.8	0.0	0.0	0.0	0.0	0.0	0.0
$\Sigma\, x_h$	515.9	772.0	1291.1	1762.1	1866.9	1964.9	1651.5	1551.0	1121.2	591.8
$\Sigma\, l_h^2$	..	24081.53	38451.92	67814.37	96932.51	107379.15	94189.71	92650.08	69923.96	38968.52
$\Sigma\, c_h^2$	20486.13	9074.71	21142.67	18535.84	0.00	0.00	0.00	0.00	0.00	0.00
$\Sigma\, x_h^2$	20486.13	33156.24	59594.59	86350.21	96932.51	107379.15	94189.71	92650.08	69923.96	38968.52
$\Sigma\, l'_{h-1}$	..	515.9	772.0	1291.1	1762.1	1866.9	1575.0	1484.3	1069.7	561.4
$\Sigma\, l'^2_{h-1}$	..	20486.13	33156.24	59594.59	86350.21	96932.51	85650.06	84860.43	63631.79	35059.24
$\Sigma\, l'_{h-1} l_h$	..	22209.57	35705.27	63570.31	91487.08	102021.16	89816.88	88667.58	66700.10	36959.62

UNIVERSITY OF MICHIGAN LIBRARIES

TABLE 5

Statistics on upper limb length in girls calculated from Table 4; data from Meredith (35)

STATISTIC	AGE IN YEARS 3	4	5	6	7	8	9	10	11	1	
L_k	0	13	18	28	36	36	29	26	18	9	
C_k	13	5	10	8	0	0	0	0	0	0	
N_k	13	18	28	36	36	36	29	26	18	9	
λ	0	.722	.643	.778	1.000	1.000	1.000	1.000	1.000	1.000	
l_k	..	43.0231	46.1889	49.1893	51.8583	54.5806	56.9483	59.6538	62.2889	65.5756	
c_k	39.6846	42.5000	45.9700	48.1000	..	..	..	..	..	..	
m_k	39.6846	42.8889	46.1107	48.9472	51.8583	54.5806	56.9483	59.6538	62.2889	65.5756	
l'_{k-1}	..	39.6846	42.8889	46.1107	48.9472	51.8583	54.3103	57.0885	59.4278	62.3778	
$\Sigma (l_k - l_k)^2$	..	18.7231	50.4778	65.9668	118.1875	133.8164	139.6325	126.9649	85.6578	54.3823	Overall Sum 793.8091
$\Sigma (c_k - c_k)^2$	12.8369	26.4520	10.2610	26.9600	..	..	..	..	..	..	..
$\Sigma (x_k - m_k)^2$	12.8369	46.0178	61.0468	100.3098	118.1875	133.8164	139.6325	126.9649	85.6578	54.3823	..
$\Sigma (l'_{k-1} - l'_{k-1})^2$	..	12.8369	46.0178	61.0468	100.3098	118.1875	111.2669	124.0266	61.8962	40.3556	675.9441
$\Sigma (l'_{k-1} - l'_{k-1})(l_k - l_k)$	..	13.9646	47.4478	62.0233	107.5109	124.7209	123.3456	123.3762	69.6756	44.4512	716.5161
$V (l_k)$	..	1.5603	2.9693	2.4432	3.3768	3.8233	4.9869	5.0786	5.0387	6.7978	..
$V (c_k)$	1.0697	6.6130	1.1401	3.8514	..	..	..	..	..	..	..
v_k	1.0697	2.7069	2.2610	2.8660	3.3768	3.8233	4.9869	5.0786	5.0387	6.7978	..
$V (l'_{k-1})$	..	1.0697	2.7069	2.2610	2.8660	3.3768	3.9738	4.9611	3.6409	5.0444	..
$cov (l'_{k-1}, l_k)$	..	1.1637	2.7910	2.2972	3.0718	3.5635	4.4052	4.9350	4.0986	5.5564	..
$b_{k,\,k-1}$	..	1.088	1.031	1.016	1.072	1.055	1.109	0.995	1.126	1.102	Overall r ..
$r_{k,\,k-1}$	..	.901	.985	.978	.987	.992	.990	.983	.957	.949	.9781
$1 - \phi_k$	0.0000	.7222	.6429	.7778	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	Equation C
ϕ_k	1.0000	.2778	.3571	.2222	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	..
M_k	39.6846	42.8778	46.1043	48.9417	51.8525	54.5746	57.2438	59.8071	62.6584	65.8945	Equation B
$_{k+1}M_k$	39.5953	42.7980	45.8955	48.9417	51.8525	54.5746	57.2438	59.8071	62.6584	65.8945	Equation F
$M_{k-k}M_{k-1}$	..	3.2825	3.3063	3.0462	2.9108	2.7221	2.6692	2.5633	2.8513	3.2361	..
$V (M_k)$	.0823	.1504	.0807	.0796	.0938	.1062	.1400	.1448	.1496	.2258	Equation H or H'
$V (_{k+1}M_k)$	.0823	.0990	.0635	.0796	.0938	.1062	.1400	.1448	.1496	.2258	Equation I or I'
$V (M_{k-k}M_{k-1})$	..	.0477	.0069	.0048	.0044	.0048	.0088	.0085	.0123	.0354	Equation P or P'
$l_k - l'_{k-1}$	..	3.3385	3.3000	3.0786	2.9111	2.7223	2.6380	2.5653	2.8611	3.1987	..
$V (l_k - l'_{k-1})$	..	.0233	.0052	.0039	.0028	.0020	.0052	.0065	.0268	.0810	Equation R
Mincr 1, 2	..	..	3.3006	3.0724	..	..	..	..	..	..	Equation N
$V (M$incr 1, 2)	..	..	.0051	.0039	..	..	..	..	..	..	Equation O

figure 384.8 for Σc_h is the sum of the age 6 values of children who were not present at age 5. The quantity Σx_h refers to *all* children at time h, so that the relation

$$\Sigma x_h = \Sigma l_h + \Sigma c_h$$

provides a check on the arithmetic. The same remarks apply to Σx_h^2, etc. (The Σc_h^2 values are not needed subsequently for any of our main purposes and could be omitted from the table; we have retained them in the general layout, however, as they may sometimes be needed, and it is of some interest to compare the variances estimated from the cross-sectional and longitudinal elements.)

The next block introduces the only possibly confusing element in the table. This refers to the children present at t_{h-1} who were also present at t_h. That is, in terms of the age column t_h, the l_h (unprimed) values throw back, the l'_h (primed) values throw forward. For computional purposes, it is easier, and more logical, to put $\Sigma l'_{h-1}$, etc., in the t_h column than in the t_{h-1} column. Thus the figure 772.0 is the sum of the values at age 4 of those children present at both ages 5 and 4. The sums of squares, $\Sigma l'^2_{h-1}$ are similarly treated.

The last row is needed in getting the correlation of the values at t_{h-1} and t_h. It is simply the sum of products of the l'_{h-1} and l_h values.

In Table 5, the first block consists of the numbers of children, L_h being the number present at t_h who were also present at t_{h-1}, C_h the number present at t_h who were absent at t_{h-1}, and N_h the total at t_h. The longitudinality fraction, λ, is calculated from L_h/N_h. The next block is of means, recorded to four decimals to prevent rounding errors when they are used in further calculations. Then comes a block of sums of squares of deviations from means and of cross-products of deviations from means. These are calculated from the sums and sums of squares of Table 4 by the usual methods. The first figure in the $\Sigma(l_h - \bar{l}_h)^2$ row, at age 4, is obtained as

$$24081.53 - (559.3)^2/13 = 18.7231.$$

The calculation of $\Sigma(c_h - \bar{c}_h)^2$ may be omitted unless the cross-sectional element variances are required for some reason such as the computation of $M_{\text{incr}\,1,2}$.

At the end of the rows for $\Sigma(l_h - \bar{l}_h)^2$, etc., at the right-hand end of the block, are recorded the overall sums of these rows. These are used in computing the overall r; thus

$$\text{overall } r = \frac{716.5161}{\sqrt{675.9441 \times 793.8091}} = 0.9781.$$

The next block records the variances and covariances, obtained from the figures in the block above by simple division (using $L_h - 1$, etc.); thus $V(l_4)$, for example, is $18.7231/(13-1) = 1.5603$. The only statistic in this block actually needed in the most usual further computations is $V(l_h)$, and the others may be omitted if desired.

The next block consists of two rows only; the upper records the regression coefficient for each set of values at t_h on the corresponding values at t_{h-1}, and the lower the corresponding correlation coefficients. Both should routinely be calculated since only by looking at these figures can one decide if the conditions of equal r for each $t_{h,h-1}$ pair is met in the data in hand. Note however that the figure for "overall r" is not obtained by averaging the r's given here, but by the process described above. The formulae for the regression and correlation coefficients are the usual ones; the regression of age 4 on age 3 values is $13.9646/12.8369 = 1.088$, and the correlation between the two, $13.9646/\sqrt{12.8369 \times 18.7231} = .901$. It will be seen that in the present example the correlation between ages 3 and 4 is rather low, and that the correlation drops also as puberty is approached at the other end of the series; but the effect is probably insufficient to invalidate the further computations, in all of which the overall r of .9781 is used as an estimate of the population figure.

The penultimate block contains the statistics we chiefly require to know. The first to be computed are $(1-\phi_h)$, and ϕ_h taken together, the former being first computed from equation C, that is

$$1 - \phi_h = \frac{L_h}{N_h - r^2 C_h \left(1 - \frac{\phi_{h-1}}{C_{h-1}} L_h\right)}, \quad C_{h-1} \neq 0.$$

For age 3, $L_h = 0$ and $(1-\phi_h) = 0$, and $\phi_h = 1$. For age 4 we have

$$(1-\phi_h) = \frac{13}{18 - .9781^2 \times 5\left(1 - \frac{1 \times 13}{13}\right)} = .7222;$$

$$\phi_h = .2778.$$

For age 5 we have

$$(1-\phi_h) = \frac{18}{28 - .9781^2 \times 10\left(1 - \frac{.2778 \times 18}{5}\right)} = .6429;$$

$$\phi_h = .3571.$$

For ages 7 to 12 we have $C_h = 0$ and $(1 - \phi_h) = 1$; $\phi_h = 0$. Note that when both $C_{h-1} = 0$ and $C_h = 0$ then $(1 - \phi_h) = 1$ and the computational complexity of equations C′ etc. does not arise; it only occurs when $C_{h-1} = 0$ and $C_h \neq 0$. A case of this occurs in Example 2 at age 6 and is illustrated below.

The successive values of M_h are now found from equation B, that is

$$M_h = (1 - \phi_h)\{\bar{l}_h + \frac{rs_h}{s_{h-1}}(M_{h-1} - \bar{l}_{h-1})\} + \phi_h \bar{c}_h .$$

At age 5 for instance, we have

$$M_h = .6429\{46.1889 + .9781 \frac{\sqrt{2.2610}}{\sqrt{2.7069}}(42.8778 - 42.8889)\} + .3571 \times 45.9700,$$

$$= 46.1043.$$

Note that $\bar{l}'_{h-1}$ is in the same column of Table 5 as ϕ_h, $\bar{c}_h$, M_h, etc., not the preceding one. Very particular care must be taken in calculating and checking the values of $(1 - \phi_h)$ and of M_h, as any error upsets all succeeding values; it is best to compute and check twice in sets of not more than five values at a time. The next row ${}_{h+1}M_h$, is now taken, using equation F, that is

$${}_{h+1}M_h = M_h - \frac{rs_h C_{h+1}\phi_h}{s_{h+1}C_h}\{M_{h+1} - \bar{c}_{h+1}\}, \quad C_h \neq 0.$$

At age 5 this time we have

$${}_{h+1}M_h = 46.1043 - .9781 \frac{\sqrt{2.2610}}{\sqrt{2.8860}} \times \frac{8}{10} \times .3571\{48.9417 - 48.1000\}$$

$$= 45.8955.$$

Note again that though at age 7 and over we have $C_h = 0$, we do not have to use equation F′ to evaluate ${}_{h+1}M_h$, as $C_{h+1} = 0$ also and thus the second term disappears, leaving ${}_{h+1}M_h = M_h$. For an instance when $C_h = 0$ and $C_{h+1} \neq 0$, see Example 2.

The best estimate of mean increment that can be obtained by these procedures, $M_h - {}_hM_{h-1}$, is then written down; note that the 4 to 5 year increment in this terminology is recorded under age 5 ($h = 5$ years) and the subtractions run diagonally, i. e.

$$\text{Increment 4 to 5 years} = 46.1043 - 42.7980$$

$$= 3.3063.$$

The last three rows of the block contain the variances of these estimates of mean and mean increment. The variance of M_h is obtained from equation H, if $C_h \neq 0$, that is

$$V(M_h) = v_h\phi_h/C_h .$$

$V(M_h)$ at age 4 is thus $(2.7069 \times .2778)/5 = .1504$.

At age 7 however $C_h = 0$ and we have to have recourse to equation H′, that is

$$V(M_h) = v_h \left\{ \frac{1 - r^2}{N_h} + \frac{r^2 V(M_{h-1})}{v_{h-1}} \right\},$$

so that at this age

$$\begin{aligned} V(M_h) &= 3.3768 \left\{ \frac{1 - .9781^2}{36} + \frac{.9781^2 \times .0796}{2.8660} \right\} \\ &= 3.3768 \times .027775 \text{ (this figure being } V(M_{h-1})/v_{h-1}) \\ &= .0938. \end{aligned}$$

The variance for ages 8 to 12 are calculated similarly for each $V(M_h)$, the $V(M_{h-1})/v_{h-1}$ having been recorded in the previous step.

The $V({}_{h+1}M_h)$ can now be recorded, if it is required. Equation I is used, that is

$$V({}_{h+1}M_h) = \frac{v_h\phi_h}{C_h} \left\{ 1 - \frac{r^2 C_{h+1}\phi_h(1 - \phi_{h+1})}{C_h} \right\}.$$

At age 4, this is

$$\begin{aligned} V({}_{h+1}M_h) &= .1504\{1 - .9781^2 \times 10/5 \times .6429 \times .2778\} \\ &= .0990. \end{aligned}$$

At age 3, $(1 - \phi_h) = 0$ and thus $V({}_{h+1}M_h) = V(M_h)$, and at age 6 and beyond, $C_{h+1} = 0$ and the same is true. It is clear, however, that the estimate ${}_{h+1}M_h$ at ages 4 and 5 is better than the estimate M_h.

Lastly the $V(M_h - {}_hM_{h-1})$ is recorded, from equation P′, in the present instance, that is

$$V(M_h - {}_hM_{h-1}) = V(M_h) + V({}_hM_{h-1}) - 2r\frac{s_h}{s_{h-1}}(1 - \phi_h)V(M_{h-1}).$$

At $h = 6$, the figures for the variance of the 5 to 6 increment are

$$\begin{aligned} V(M_h - {}_hM_{h-1}) &= .0796 + .0635 - 2 \times .9781 \times \frac{\sqrt{2.8660}}{\sqrt{2.2610}} \\ &\quad \times .7778 \times .0807 \\ &= .0048. \end{aligned}$$

In the last block are recorded for comparative purposes the longitudinal element mean increment $\bar{l}_h - \bar{l}'_{h-1}$, and its variance from equation R, that is

$$V(\bar{l}_h - \bar{l}'_{h-1}) = \{V(l_h) + V(l'_{h-1}) - 2\,\mathrm{cov}(l_h, l'_{h-1})\}/L_h,$$

and the mean increment estimated from the present and previous occasion only, $M_{\mathrm{incr}\,1,2}$ from equation N that is

$$M_{\mathrm{incr}\,1,2} = \frac{A(\bar{l}_h - \bar{l}'_{h-1}) + B(\bar{c}_h - \bar{c}'_{h-1})}{A + B},$$

with

$$A = \frac{V(c_h)}{C_h} + \frac{V(c'_{h-1})}{C'_{h-1}},$$

$$B = \frac{V(l_h) + V(l'_{h-1}) - 2\,\mathrm{cov}(l_h, l'_{h-1})}{L_h},$$

and its variance, from equation O, that is

$$V(M_{\mathrm{incr}\,1,2}) = \frac{AB}{A + B}.$$

The computations for the age 5 to 6 increments are as follows

$$\bar{l}_h - \bar{l}'_{h-1} = 49.1893 - 46.1107$$
$$= 3.0786,$$

$$V(\bar{l}_h - \bar{l}'_{h-1}) = \frac{2.4432 + 2.2610 - 2 \times 2.2972}{28}$$
$$= .0039,$$

$$M_{\mathrm{incr}\,1,2} = \frac{\left\{\frac{3.8514}{8} + \frac{1.1401}{10}\right\}(49.1893 - 46.1107) + \left\{\frac{2.4432 + 2.2610 - 2\times 2.2972}{28}\right\}(48.1000 - 45.9700)}{\left\{\frac{3.8514}{8} + \frac{1.1401}{10}\right\} + \left\{\frac{2.4432 + 2.2610 - 2\times 2.2972}{28}\right\}}$$
$$= 3.0724.$$

The content of the second curly bracket is $V(\bar{l}_h - \bar{l}'_{h-1})$, calculated already, but this will not usually be directly available under the ordinary circumstances in which $M_{\mathrm{incr}\,1,2}$ would be used. Lastly

$$V(M_{\mathrm{incr}\,1,2}) = \frac{(10\times 3.8514 + 8\times 1.4011)(2.4432 + 2.2610 - 2\times 2.2972)}{28(10\times 3.8514 + 8\times 1.4011) + 10\times 8(2.4432 + 2.2610 - 2\times 2.2972)}$$
$$= .0039.$$

Example 2. The data from Maresh's study are reproduced in Table 6, and in Table 7 the various statistics are displayed. The layout

TABLE 6

Length of humerus (cms. measured by x-ray) from 6 months to 12 years in healthy girls; data from Maresh (36)

SUBJECT NUMBER	AGE IN YEARS ½	1	1½	2	2½	3	3½	4	4½	5	5½	6
1	8.1	..	..	..	..	..	..	..	..	..	..	..
2	9.0	..	..	..	..	..	..	..	..	..	..	..
3	8.4	..	..	..	..	..	..	..	..	..	..	..
5	..	11.1	12.3	13.5	..	..	..	..	..	..	..	..
6	..	9.7	11.1	..	..	..	..	..	..	..	..	..
7	9.1	10.6	12.2	13.6	..	..	..	..	..	..	..	..
8	9.0	10.7	12.3	13.1	..	..	..	..	..	..	..	..
9	9.5	11.0	12.7	..	..	..	..	..	..	..	..	..
10	8.9	10.5	12.0	..	..	..	..	..	..	..	..	..
11	8.8	10.5	11.6	12.9	..	..	..	..	..	..	..	..
12	9.0	10.9	..	..	..	..	..	..	..	..	..	..
13	9.0	10.6	11.2	12.4	13.1	13.9	..	..	..	..	..	..
14	..	..	11.8	13.2	..	..	..	..	..	..	..	..
15	..	10.5	11.8	12.9	13.8	14.6	..	..	..	..	..	..
16	..	10.3	11.6	12.6	13.8	14.3	..	..	..	..	..	..
17	..	10.4	11.6	12.4	13.5	14.1	14.9	..	..	..	..	..
18	..	..	10.9	11.5	12.7	13.6	14.3	15.0	..	..	..	..
19	..	10.4	11.7	12.9	13.8	14.5	15.5	..	..	..	..	..
20	..	10.0	11.2	12.1	13.0	13.8	14.4	15.2	16.0	..	..	..
21	..	..	..	12.8	13.6	14.3	15.3	16.1	16.5	..	..	..
22	..	..	11.8	13.6	14.7	15.7	16.4	17.0	18.0	..	..	..
23	..	..	12.4	..	14.8	15.6	16.7	17.5	18.2	..	..	..
24	..	..	..	12.8	14.1	15.2	16.0	16.5	17.6	..	..	..
25	..	..	..	..	13.2	..	15.8	16.4	17.3	..	18.8	..
26	..	..	..	..	12.9	13.3	14.1	14.8	15.5	16.2	16.9	..
27	..	..	..	..	..	..	..	..	..	19.2	19.9	20.6
28	..	..	..	..	..	15.6	..	17.4	18.3	18.8	19.9	20.7
29	..	..	..	..	..	..	16.7	17.4	18.5	19.2	20.1	..
30	..	..	..	..	..	..	..	..	..	19.5	..	..
31	..	..	..	..	..	..	..	..	..	16.4	17.0	17.6
32	..	..	..	..	..	..	..	..	..	..	..	..
33	..	..	..	..	..	..	..	..	16.2	17.0	17.7	18.3
34	..	..	..	..	..	..	..	..	..	17.8	..	..
$\Sigma\, l_k$	0.0	74.8	153.3	166.7	136.1	172.9	137.6	145.9	155.9	71.2	111.5	77.2
$\Sigma\, c_k$	88.8	72.4	46.9	25.6	40.9	15.6	32.5	17.4	16.2	72.9	18.8	0.0
$\Sigma\, x_k$	88.8	147.2	200.2	192.3	177.0	188.5	170.1	163.3	172.1	144.1	130.3	77.2
$\Sigma\, l_k^2$	0.00	799.52	1810.61	2142.19	1855.33	2497.59	2110.86	2373.51	2709.93	1273.52	2083.93	1497.50
$\Sigma\, c_k^2$	789.88	749.96	551.05	327.68	559.69	243.36	528.53	302.76	262.44	1334.69	353.44	0.00
$\Sigma\, x_k^2$	789.88	1549.48	2361.66	2469.87	2415.02	2740.95	2639.39	2676.27	2972.37	2608.21	2437.37	1497.50
$\Sigma\, l'_{k-1}$	0.0	63.3	136.3	152.0	126.0	163.8	130.1	139.7	148.3	68.5	106.8	74.5
$\Sigma\, l'^2_{k-1}$	0.00	572.71	1430.67	1779.40	1590.40	2240.78	1886.73	2177.13	2451.27	1179.83	1911.12	1394.31
$\Sigma\, l'_{k-1} l_k$	0.00	676.61	1609.05	1951.88	1717.59	2365.52	1995.56	2273.14	2577.19	1225.74	1995.60	1444.98

TABLE 6 (Continued)

Length of humerus (cms. measured by x-ray) from 6 months to 12 years in healthy girls; data from Maresh (36)

SUBJECT NUMBER	AGE IN YEARS 6½	7	7½	8	8½	9	9½	10	10½	11	11½	12
27	21.3	22.0	..	..	..	..	..	..	..	..	..	..
28	21.1	..	..	..	..	..	..	..	..	..	..	..
29	..	..	..	..	..	..	..	..	..	..	..	..
30	22.2	23.3	23.1	24.4	..	..	..	..	..	..	..	..
31	18.0	18.8	19.0	..	..	..	..	..	..	..	..	..
32	20.6	21.2	21.8	22.6	23.4	24.1	24.7	25.2	..	..	..	..
33	..	..	..	..	..	..	..	..	..	..	..	..
34	19.8	20.0	21.0	21.7	22.3	23.2	24.1	..	..	..	..	..
35	20.3	20.4	..	22.1	22.8	23.0	23.8	..	..	..	..	..
36	..	..	21.3	..	22.5	23.3	23.6	24.2	..	..	..	..
37	19.8	..	..	21.5	22.1	22.9	23.4	24.1	..	..	..	..
38	..	19.3	19.7	20.6	21.6	22.4	23.0	23.3	..	..	..	..
39	19.2	..	20.2	20.9	21.3	22.2	22.7	..	..	..	..	..
40	..	..	..	23.1	..	24.1	24.8	25.3	26.5	..	..	..
41	20.7	..	..	..	23.6	24.5	25.2	25.8	26.3	27.4	..	..
42	..	..	20.7	21.3	21.8	22.3	23.2	23.9	..	..	..	..
43	..	..	21.6	21.9	22.9	23.5	24.4	24.7	25.4	..	..	..
44	..	..	21.8	22.6	..	..	24.7	25.0	25.6	26.5	..	..
45	..	19.8	..	21.3	21.8	22.5	..	23.7	..	..	..	..
46	..	..	..	22.3	..	23.4	..	..	..	..	..	..
47	..	..	..	..	..	..	25.8	..	..	..	29.5	30.0
48	..	..	..	..	..	21.2	..	21.7	22.4	22.8	23.4	23.9
49	..	..	..	..	21.5	..	..	..	..	..	..	..
50	..	..	..	..	..	..	..	26.5	..	..	..	..
51	..	..	..	..	..	..	..	25.1	25.8	..	27.2	28.1
52	..	..	..	..	..	..	..	..	25.3	25.9	26.9	27.6
$\Sigma\, l_h$	60.4	125.7	104.6	176.0	200.0	253.9	262.9	221.5	152.0	102.6	50.3	109.6
$\Sigma\, c_h$	142.6	39.1	105.6	110.3	67.6	68.7	50.5	97.0	25.3	0.0	56.7	0.0
$\Sigma\, x_h$	203.0	164.8	210.2	286.3	267.6	322.6	313.4	318.5	177.3	102.6	107.0	109.6
$\Sigma\, l_h^2$	1222.90	2645.93	2198.94	3882.24	4448.24	5865.99	6289.83	5456.41	3861.86	2643.66	1271.17	3022.58
$\Sigma\, c_h^2$	2910.50	764.53	2232.02	2435.25	1525.46	1577.81	1275.73	2364.84	640.09	0.00	1610.09	0.00
$\Sigma\, x_h^2$	4133.40	3410.46	4430.96	6317.49	5973.70	7443.80	7565.56	7821.25	4501.95	2643.66	2881.26	3022.58
$\Sigma\, l'_{h-1}$	58.9	122.2	102.6	169.9	193.9	246.1	255.5	217.0	147.6	99.6	48.7	107.0
$\Sigma\, l'^2_{h-1}$	1162.61	2499.02	2118.26	3616.27	4180.47	5511.45	5940.55	5237.38	3641.72	2488.90	1190.65	2881.26
$\Sigma\, l'_{h-1} l_h$	1192.35	2571.10	2157.80	3746.65	4312.10	5685.73	6112.51	5345.66	3750.03	2565.01	1230.23	2951.02

UNIVERSITY OF MICHIGAN LIBRARIES

TABLE 7

Statistics on length of humerus in girls calculated from Table 6; data from Maresh (36)

	AGE IN YEARS											
	½	1	1½	2	2½	3	3½	4	4½	5	5½	6
L_k	0	7	13	13	10	12	9	9	9	4	6	4
C_k	10	7	4	2	3	1	2	1	1	4	1	0
N_k	10	14	17	15	13	13	11	10	10	8	7	4
λ	0	.500	.767	.867	.769	.923	.818	.900	.900	.500	.857	1.000
l_k	..	10.6857	11.7923	12.8231	13.6100	14.4083	15.2889	16.2111	17.3222	17.8000	18.5833	19.3000
c_k	8.8800	10.3429	11.7250	12.8000	13.6333	15.6000	16.2500	17.4000	16.2000	18.2250	18.8000	..
m_k	8.8800	10.5143	11.7765	12.8200	13.6154	14.5000	15.4636	16.3300	17.2100	18.0125	18.6143	19.3000
l_{k-1}	..	9.0429	10.4846	11.6923	12.6000	13.6500	14.4456	15.5222	16.4778	17.1250	17.8000	18.6250
$V\ (l_k)$	..	.0381	.2374	.3819	.3343	.5808	.8886	1.0386	1.1744	2.0533	2.3777	2.5133
$V\ (c_k)$	.1484	.1895	.3825	..	1.0433	..	.3950	..	..	2.0292	..	..
c_k	.1484	.1367	.2507	.3274	.4247	.6417	.9025	1.0645	1.1699	1.8013	1.9883	2.5133
$V\ (l'_{k-1})$	..	.0495	.1347	.1808	.3111	.4464	.7578	1.0844	.9519	2.2558	2.0160	2.2492
$cov\ (l'_{k-1}, l_k)$	..	.0340	.1465	.2310	.3033	.4941	.8094	1.0560	1.0381	2.1467	2.1800	2.3767
$b_{k,\,k-1}$	..	.688	1.088	1.278	.975	1.107	1.068	.974	1.090	.952	1.081	1.057
$r_{k,\,k-1}$	..	.784	.819	.879	.940	.970	.986	.995	.981	.998	.996	1.000
$1 - \phi_k$	0.0000	.5835	.8057	.9093	.8745	.9581	.9173	.9574	.9563	.8247	.9529	1.0000
ϕ_k	1.0000	.4165	.1943	.0907	.1255	.0419	.0827	.0426	.0437	.1753	.0471	0.0000
M_k	8.8800	10.4538	11.7465	12.8761	13.8817	14.7247	15.6651	16.4068	17.2036	17.9530	18.6888	9.3701
$_{k+1}M_k$	8.8010	10.4501	11.7402	12.8471	13.9108	14.7650	15.7020	16.3670	17.2411	17.9575	18.6888	9.5200
$M_k - {}_kM_{k-1}$	..	1.6528	1.2964	1.1359	1.0346	.8139	.9001	.7048	.8366	.7119	.7313	.6813
$V\ (M_k)$	.0148	.0081	.0122	.0148	.0178	.0269	.0373	.0454	.0511	.0789	.0934	.1415
$V\ ({}_{k+1}M_k)$	.0090	.0066	.0112	.0131	.0171	.0249	.0359	.0436	.0441	.0758	.0934	.1056
$V\ (M_k - {}_kM_{k-1})$	..	.0009	.0015	.0012	.0021	.0030	.0050	.0055	.0058	.0209	.0149	.0298
$l_k - l'_{k-1}$	..	1.6428	1.3077	1.1308	1.0100	.7583	.8433	.6889	.8444	.6750	.7833	.6750
$V\ (l_k - l'_{k-1})$	..	.0028	.0061	.0078	.0039	.0033	.0031	.0012	.0055	.0038	.0057	.0015

TABLE 7 (Continued)

Statistics on length of humerus in girls calculated from Table 6; data from Maresh (36)

	AGE IN YEARS											
	6½	7	7½	8	8½	9	9½	10	10½	11	11½	12
L_k	3	6	5	8	9	11	11	9	6	4	2	4
C_k	7	2	5	5	3	3	2	4	1	0	2	0
N_k	10	8	10	13	12	14	13	13	7	4	4	4
λ	.300	.750	.500	.615	.750	.786	.846	.692	.857	1.000	.500	1.000
l_k	20.1333	20.9500	20.9200	22.0000	22.2222	23.0818	23.9000	24.6111	25.3333	25.6500	25.1500	27.4000
c_k	20.3714	19.5500	21.1200	22.0600	22.5333	22.9000	25.2500	24.2500	25.3000	...	28.3500	...
m_k	20.3000	20.6000	21.0200	22.0231	22.3000	23.0429	24.1077	24.5000	25.3286	25.6500	26.7500	27.4000
l'_{k-1}	19.6333	20.3667	20.5200	21.2375	21.5444	22.3727	23.2273	24.1111	24.6000	24.9000	24.3500	26.7500
$V(l_k)$	3.4233	2.5030	2.6770	1.4629	.4744	.5516	.6520	.6311	2.2387	3.9900	6.1250	6.5133
$V(c_k)$	.9224	.1250	.4370	.5080	1.1033	2.2900	.6050	4.1966	...	...	2.6450	...
v_k	1.3889	2.2257	1.3951	1.0236	.5655	.7826	.8508	1.5000	1.8657	3.9900	6.3367	6.5133
$V(l'_{k-1})$	3.1033	2.0427	3.2270	1.1455	.3753	.5522	.5982	.6586	2.1520	2.9533	4.8050	6.3367
$cov(l'_{k-1}, l_k)$	3.2483	2.2020	2.8520	1.2643	.4014	.5295	.6060	.6311	2.1660	3.4233	5.4250	6.4067
$b_{k,\,k-1}$	1.047	1.078	.884	1.103	1.070	.959	1.013	.958	1.007	1.159	1.129	1.011
$r_{k,\,k-1}$	.996	.974	.971	.976	.951	.959	.970	.979	.987	.997	1.000	.998
												Overall r .9769
$1 - \phi_k$	.6744	.9058	.7871	.8119	.8905	.8953	.9303	.8670	.9621	1.0000	.8799	1.0000
ϕ_k	.3256	.0942	.2129	.1881	.1095	.1047	.0697	.1330	.0379	0.0000	.1201	0.0000
M_k	20.0818	20.4991	20.9498	21.8158	22.4317	23.1234	23.8956	24.3207	25.0393	25.8489	27.1582	27.8043
$_{k+1}M_k$	20.0137	20.5486	21.0091	21.8309	22.4114	23.2120	23.8884	24.3283	25.0393	25.9369	27.1582	27.8043
$M_k - {}_kM_{k-1}$	.5618	.4854	.4012	.8067	.6008	.7120	.6836	.4323	.7110	.8096	1.2213	.6461
$V(M_k)$	.0646	.1048	.0594	.0385	.0206	.0273	.0297	.0499	.0717	.1899	.3805	.4420
$V({}_{k+1}M_k)$	.0594	.0863	.0496	.0348	.0187	.0256	.0263	.0484	.0717	.1747	.3805	.4420
$V(M_k - {}_kM_{k-1})$	.0315	.0195	.0181	.0074	.0056	.0036	.0036	.0094	.0155	.0568	.1437	.0687
$l_k - l'_{k-1}$	.5000	.5833	.4000	.7625	.6778	.7091	.6727	.5000	.7333	.7500	.9000	.6500
$V(l_k - l'_{k-1})$	.0100	.0237	.0400	.0101	.0052	.0041	.0035	.0031	.0098	.0243	.0400	.0095

follows that of the previous example, except that the sums of the squared deviations have been omitted for economy of space and $M_{\text{incr}\,1,2}$ and its variance are not given. The computations are similar to those in Example 1 and need not be repeated here. At age $6\frac{1}{2}$, however, there is an example of the calculation of $1-\phi_h$ when $C_{h-1}=0$ and $C_h \neq 0$. The computation proceeds as follows:

Since $C_{h-1}=0$, we must use equations C', C'' and C'''. We find first

$$\frac{V(M_{h-2})}{v_{h-2}} = \frac{\phi_{h-2}}{C_{h-2}} \qquad C'''$$

$$= \frac{.0469}{1}$$

$$= .0469.$$

Then substituting this value in equation C'' we have

$$\frac{V(M_{h-1})}{v_{h-1}} = \frac{1-r^2}{N_{h-1}} + r^2 \frac{V(M_{h-2})}{v_{h-2}} \qquad C''$$

$$= \frac{1-.9769^2}{4} + .9769^2 \times .0469$$

$$= .0562.$$

Finally, substituting this value in equation C'

$$1-\phi_h = \frac{L_h}{N_h - r^2 C_h\{1 - L_h V(M_{h-1})/v_{h-1}\}} \qquad C'$$

we have

$$1-\phi_h = \frac{3}{10 - .9769^2 \times 7(1 - 3 \times .0562)}$$

$$= .6747.$$

A similar calculation must be done to find ϕ_h at age $11\frac{1}{2}$ since C_h at 11 is zero.

The same difficulty arises in the computation of ${}_{h+1}M_h$ at age 6, when $C_h = 0$. We have then from equation F

$${}_{h+1}M_h = 19.3701 - .9769 \times \frac{\sqrt{2.5133}}{\sqrt{1.9883}} \times 7 \times \frac{V(M_h)}{v_h}\{20.0818 - 20.3714\},$$

and by equation C″

$$\frac{V(M_h)}{v_h} = \frac{1 - .9769^2}{4} + .9769^2 \frac{V(M_{h-1})}{v_{h-1}},$$

and by equation C‴

$$\frac{V(M_{h-1})}{v_{h-1}} = \frac{\phi_{h-1}}{C_{h-1}} = \frac{.0471}{1} = .0471.$$

Whence $\frac{V(M_h)}{v_h} = .0563$ and ${}_{h+1}M_h = 19.5200$.

UNIVERSITY OF MICHIGAN LIBRARIES

METHODOLOGY IN THE STUDY OF PHYSICAL MEASUREMENTS OF SCHOOL CHILDREN*

PART I

BY B. G. GREENBERG AND A. HUGHES BRYAN

University of North Carolina, Chapel Hill

IN PREVIOUS studies by the authors (3, 4), it was recommended that the technique of covariance be employed in the analysis of anthropometric characteristics of school children. It obviates the classical requirement for balancing large homogeneous groups of children according to age, height, or weight in order to ascertain whether or not observed differences in another variable are random fluctuations. The advantages of covariance in this type of application have also been pointed out recently by Kemsley (14) in a study of weight and height of a population. In the present study, it can be shown that the effect of adjustment for age by covariance was equivalent, at least, to doubling the number of observations.

The ultimate aim of the continuing series of investigations is to examine the etiology, in terms of nutritional status, for physical differences in school children of differing socio-economic backgrounds. In pursuing this objective, several problems in methodology have had to be resolved to assure maximum efficiency at each step. The purpose of the present paper is to report some of these results. They include the mechanics of the sampling used to obtain physical measurements, the optimal age group for study, and the physical measurements which appear to be worthwhile indices to adopt.

a. Sampling technique used in obtaining measurements

The following measurements were made on boys and girls: age (months), weight (pounds), height (centimeters), transverse chest dia-

* A preliminary report of this paper was presented before a joint meeting of the Biometric Society and the Institute of Mathematical Statistics, March 17, 1950, Chapel Hill, North Carolina.

meter at the level of the xiphoid (centimeters), bi-iliac diameter (centimeters), and calf circumference (centimeters). Calf circumference may be interpreted as an indicator of the amount of muscle and subcutaneous fat, whereas the chest and bi-iliac diameters are gauges of skeletal stockiness.

It was expected that there would be variability in these measurements due to the children selected, the observers (both within and between), and whether the children were dressed or undressed. A series of pilot studies were undertaken in which each of several trained observers made duplicate but independent measurements on a number of children. In certain of these studies, the children were first dressed and then undressed. Estimation of the separate components of variance was then made from each of these experiments. The results are listed briefly herewith.

Except for the boys' bi-iliac measurements, clothes made no significant change in the size of the differences between two group means. For example, one group of ten boys when dressed measured 0.8 centimeters more than another group in the transverse chest diameter, whereas this difference was 0.9 centimeters when the boys were undressed. The change from 0.8 to 0.9 was not significant. On the other hand, the same ten boys measured 0.7 centimeters more than another group in the bi-iliac diameter while wearing clothes, and this difference decreased significantly to 0.5 centimeters when the children were undressed. In this manner, it was observed that the extent of the differences between group means would be little changed if all the children were dressed or undressed, with the exception of the bi-iliac diameter. The reason that the boys' bi-iliac diameter measurements differed significantly might be explained by the variety in the types of trousers and coveralls worn at the time of measurement. (Of course, measurements must be taken at the same season of the year so that the type of garment worn will not otherwise disturb the equilibrium.)

Furthermore, the variance of the difference between group means was not significantly increased when clothes were worn. Such information was beneficial from the standpoint of both economy and avoiding the problem of convincing school authorities that children should undress. This was especially valuable in the case of girls. All subsequent measurements reported herein are based, therefore, on children who were dressed. Unfortunately, a correction factor for clothes cannot be assigned to the measurements reported in the main body of this paper, since the pilot

UNIVERSITY OF MICHIGAN LIBRARIES

studies were undertaken with a different group of children and at varying times of the year.

To minimize "observer errors" in conjunction with the number of children expected to be in the sample at a later date, it was found that an overlap in measurement to the extent of one in three was desirable. That is, observer A should measure the first child, observer B the second child, and both A and B should take independent measurements on the third child. The only exception to this plan was in weighing children, where the almost complete absence of within-observer variability in the pilot study indicated such overlap was unnecessary. A description of the pilot studies and their composition is contained in the Appendix. The recommendation for a sampling design was based upon estimates of the components of variance derived from these studies and may be found in Table 10 of the Appendix.

b. Optimal age groups for study

The ages of children studied varied from 72 to 167 months (that is, 6 years and up to but not including 14 years, which for brevity will be referred to as "6 thru 13 years"). This is a group of children ordinarily enrolled in an elementary school. Using the sampling plan in (a), two schools for white children in North Carolina from differing socio-economic backgrounds were studied. Children in school A came from urban families with presumably higher than average incomes, and those in school B came from rural families in a cash farming region. There were 292 children in A and 154 in B, and it was believed that each constituted a relatively homogeneous group with respect to socio-economic level.[1]

The problem arose as to which age grouping was optimal for showing physical differences between the children from the two schools. Too large an age interval would violate the assumption of homogeneous vari-

[1] It was assumed that genetic or racial factors would not play a major role in any physical differences such as weight and height between these two schools. These children are from a group of families indigenous to North Carolina for several generations. Furthermore, studies by Boas (2) on the relative role of heredity and environment indicate that certain traits, such as the form of the head and face are much more strongly influenced by race than are other features such as height and weight. Also, Stuart (15) has stated, "Several recent studies offer strongly suggestive evidence that dietary habits or factors associated with socio-economic circumstances may have more to do with the differences found to exist between racial groups than has the factor of race itself."

ances within the group. For example, it has been shown by Burgess (6) that the standard deviation and coefficient of variation for height are functions of age. Too great an age grouping would also result in having the variable under consideration deviate from linear regression on age. This was undesirable computationally, since it was hoped to derive a simple routine for broad application. There are numerous small sample nutritional studies where an experimental group receives, say, some class of vitamins, and the effectiveness of the therapy is measured by comparing physical measurement against a control group. It would be worthwhile as a secondary objective here to develop efficient yet simple techniques of comparison for such projects.

Optimal was defined rather loosely in terms of several considerations, viz., the coefficient of variation, the number of children who would be in that age group in the two schools, the magnitude of the differences between adjusted group means, and the probability associated with the *F*-ratio measuring the significance of those differences. It is clear that a particular age grouping might have a low coefficient of variation, but if there is no real difference between mean height, or mean weight, of children in that age group from the two schools, the *F*-ratio would probably be not signicant. On the other hand, a high *F*-ratio would not imply, *per se*, a difference between group means which would be of biological value.

The various age groupings which were tried were 72-107 months (6 thru 8 years), 108-143 months (9 thru 11 years), 132-167 months (11 thru 13 years), and 96-143 months (8 thru 11 years). It was deemed preferable to choose identical age groupings for both sexes, since school health programs are conveniently administered through the classroom. The optimal division for both boys and girls, in terms of the above considerations, was the 96-143 months age grouping.

For boys alone, 108-143 months appeared about as suitable as the larger group inclusive of the eight-year-olds. It was decided to choose the 96-143 months group, however, when consideration was given to all the measurements adjusted for both age and height. For girls alone, 72-107 months seemed slightly more desirable than 96-143 months, particularly when all the measurements were adjusted for both age and height. The latter was chosen in spite of this to retain uniformity with the boys. Furthermore, there were greater physical differences between the mean heights and mean weights in the 96-143 months group, which

UNIVERSITY OF MICHIGAN LIBRARIES

made it more favorable from the standpoint of relating growth to nutritional background.

Both boys and girls in the 96-143 months age group in school A were significantly taller and heavier for their age than those in school B of a comparable age group. These results are presented in Tables 1 and 2.

TABLE 1

Observed mean ages, heights, and weights for children from two schools in North Carolina by sex and age groups

AGE GROUPINGS	NUMBER OF CHILDREN		OBSERVED GROUP MEANS						
			Age (mos.)			*Height (cms.)*		*Weight (lbs.)*	
	A	B	A	B	Total	A	B	A	B
Boys									
72-107 mos. (6 thru 8 yrs.)	59	31	91.9	93.8	92.5	125.8	124.2	57.5	56.0
96-143 mos. (8 thru 11 yrs.)	87	47	119.3	117.1	118.5	139.4	135.6	77.2	67.0
132-167 mos. (11 thru 13 yrs.)	55	31	147.8	149.4	148.4	151.4	148.4	93.2	89.9
108-143 mos. (9 thru 11 yrs.)	66	30	124.6	125.2	124.8	142.3	138.4	81.1	69.9
Girls									
72-107 mos. (6 thru 8 yrs.)	49	20	90.5	89.3	90.2	125.0	121.8	58.0	49.6
96-143 mos. (8 thru 11 yrs.)	72	44	121.1	122.1	121.5	140.6	137.7	78.1	70.8
132-167 mos. (11 thru 13 yrs.)	46	29	145.7	144.7	145.3	152.0	148.7	95.7	89.7
108-143 mos. (9 thru 11 yrs.)	59	38	125.2	125.5	125.3	142.9	139.4	80.4	73.6

It can be seen from Table 2 that not only were the differences and the significance of the differences frequently a maximum at these ages, but the children in school B tended to diminish the magnitude of this differential as they became older, especially with respect to weight. Thus, as the two differing socio-economic groups of children grew older, the difference in mean height decreased slightly, and the difference in mean weight almost vanished. This led us to believe that the older children in school B grew heavier for their age and height than those in school A. This was tested, and the results are presented in Table 3.

TABLE 2

Group mean measurements adjusted for age

AGE GROUPINGS	HEIGHT (cms.) ADJUSTED FOR AGE[1]					WEIGHT (lbs.) ADJUSTED FOR AGE[1]				
	A	B	A-B	F-ratio of diff.	Coefficient of variation	A	B	A-B	F-ratio of diff.	Coefficient of variation
Boys										
72-107 mos. (6 thru 8 yrs.)	125.9	123.9	2.0	2.34	5.3	57.9	55.2	2.7	1.40	17.5
96-143 mos. (8 thru 11 yrs.)	139.4	135.7	3.7	11.18**	5.0	76.7	68.0	8.7	9.75**	20.8
132-167 mos. (11 thru 13 yrs.)	151.6	148.1	3.5	5.13*	5.4	90.4	89.1	1.3	0.91	23.5
108-143 mos. (9 thru 11 yrs.)	142.6	137.8	4.8	14.49**	4.6	81.3	69.5	11.8	12.27**	19.6
Girls										
72-107 mos. (6 thru 8 yrs.)	124.9	122.0	2.9	4.16*	4.8	57.7	50.2	7.5	10.91**	15.4
96-143 mos. (8 thru 11 yrs.)	140.7	137.5	3.2	6.56*	5.5	78.4	70.3	8.1	8.87**	18.9
132-167 mos. (11 thru 13 yrs.)	151.9	148.9	3.0	3.79	5.0	95.5	90.1	5.4	1.22	21.7
108-143 mos. (9 thru 11 yrs.)	142.9	139.4	3.5	6.30*	5.4	80.4	73.5	6.9	4.41*	20.4

[1] The means were adjusted for age differences indicated by columns 3, 4, and 5 in Table 1.

* Significant at 5% level.

** Highly significant at 1% level.

UNIVERSITY OF MICHIGAN LIBRARIES

Although such a trend was borne out by the data in Table 3, the differences were not significant. That being the case, the conjecture was made that there are factors other than age and height which influence the mean weight of differing socio-economic groups.

Based upon the sampling scheme in part (a) and estimates of

TABLE 3

Group mean weight adjusted for age and height

AGE GROUPINGS	WEIGHT (lbs.) ADJUSTED FOR AGE AND HEIGHT[1]				
	A	B	A-B	F-ratio of difference	Coefficient of variation
Boys					
72- 107 mos. (6 thru 8 yrs.)	57.2	56.4	0.8	0.33	11.8
96-143 mos. (8 thru 11 yrs.)	74.8	71.4	3.4	2.33	18.5
132-167 mos. (11 thru 13 yrs.)	92.1	93.5	−1.4	0.15	19.7
108-143 mos. (9 thru 11 yrs.)	79.0	74.7	4.3	2.05	18.7
Girls					
72-107 mos. (6 thru 8 yrs.)	56.6	52.2	4.4	7.54**	11.8
96-143 mos. (8 thru 11 yrs.)	76.2	73.2	3.0	1.58	18.8
132-167 mos. (11 thru 13 yrs.)	92.3	92.9	−0.6	0.02	19.7
108-143 mos. (9 thru 11 yrs.)	78.2	76.3	1.9	0.44	19.4

[1] The means were adjusted for age and height differences indicated in Table 1.
** Highly significant at 1% level.

variation from the two schools, a table was prepared to show the number of children 96-143 months old required to establish significance of differences in weight and height measurements. For example, to establish a difference significant from zero when the true difference is at least one centimeter in height, adjusted for age, would require an equal number of 378 boys in each socio-economic group, in comparison to 465 girls that would be required in each; a difference of two centimeters would decrease this number to 96 boys, or 118 girls, in each group, etc. These results are presented in Table 4, with supplementary

TABLE 4

*Approximate numbers of observations required to establish significance ($P = .05$) of differences in adjusted means as large or larger than those indicated, and also significance ($P = .05$) of true differences of stated sizes with .95 probability**

MEASUREMENT	BOYS								GIRLS							
	Same number in each school				Twice as many in one school				Same number in each school				Twice as many in one school			
	Differences (lbs.)				*Differences* (lbs.)				*Differences* (lbs.)				*Differences* (lbs.)			
	3	5	7	9	3	5	7	9	3	5	7	9	3	5	7	9
Weight (on age)	**211**	**77**	**40**	**25**	**159**	**58**	**30**	**19**	**181**	**67**	**35**	**22**	**136**	**51**	**27**	**17**
	211	**77**	**40**	**25**	**318**	**116**	**60**	**38**	**181**	**67**	**35**	**22**	**272**	**102**	**54**	**34**
	569	*205*	*105*	*65*	*426*	*153*	*78*	*48*	*489*	*177*	*91*	*56*	*366*	*132*	*68*	*42*
	569	*205*	*105*	*65*	*852*	*306*	*156*	*96*	*489*	*177*	*91*	*56*	*732*	*264*	*136*	*84*
	Differences (cms.)				*Differences* (cms.)				*Differences* (cms.)				*Differences* (cms.)			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Height (on age)	**378**	**96**	**44**	**25**	**284**	**72**	**33**	**19**	**465**	**118**	**53**	**31**	**349**	**89**	**40**	**24**
	378	**96**	**44**	**25**	**568**	**144**	**66**	**38**	**465**	**118**	**53**	**31**	**698**	**178**	**80**	**48**
	1020	*257*	*115*	*66*	*765*	*192*	*86*	*49*	*1258*	*317*	*141*	*81*	*943*	*237*	*105*	*60*
	1020	*257*	*115*	*66*	*1530*	*384*	*172*	*98*	*1258*	*317*	*141*	*81*	*1886*	*474*	*210*	*120*

* For each measurement, the superior figures in boldface type represent the number of cases required when the true difference between the two groups is as large as that indicated and the test of significance (at 5% level) is only that the two groups differ. The inferior figures in italics indicate the number of cases required to establish significance (at 5% level) of the true differences of stated size with the probability of doing so equal to 95%. These calculations were based upon estimates of variances obtained in analyzing measurements of urban and rural school children aged 96-143 months (8 thru 11 years). The estimates of variance had 131 degrees of freedom for boys, and 113 for girls. Consideration has been given to variation caused by regression on age.

UNIVERSITY OF MICHIGAN LIBRARIES

information for unequal groups in which one might be twice the size of the other. In this table, there is also presented the minimum number of observations which will, with a probability of 95 per cent establish differences of certain stated sizes to be significant at the 5 per cent level. The size of these latter groups is correspondingly larger, since an alternate hypothesis is specified with the Type II error fixed at 0.05. Tables from the paper by Harris, Horvitz, and Mood (10) were used for the purpose of checking the results.

In preparing this table, consideration was given not only to the variation in height (or weight) but also to the sampling error of the regression on age. That is, the "effective error variance" per child was used. An average, or expected, value of this effective error variance in terms of the degrees of freedom for error, suggested by Cochran (7) in 1940 and defined by Finney (9) in 1946, was found to be an excellent approximation in this case.

It should be mentioned that 96-143 months is considered a good age range by nutritionists for obtaining fairly reliable dietary information from the children themselves (5). This was a fortunate coincidence of age grouping in this instance, since the hypothesis that physical differences of growth and maturation may be explained in part by dietary patterns is to be tested in later work.[2]

[2] Since body measurements are the result of nutrition and other factors over a period of years, it would be most desirable to study dietary patterns in children prior to the manifestation of physical differences. When the diets of younger children (e.g., pre-96 months) are studied, however, the question is raised whether physical differences of growth will occur later. To answer this unequivocally, a long-term follow-up study of selected children is necessary. Under present circumstances, an involved study of this caliber was not possible.

Suppose, however, that the physical differences observed in the present instance represent a characteristic of the age grouping 96-143 months itself and that such differences will be repeated in the future as other cohorts of children reach this age. In that case, it would be sound procedure to substitute for the long-term study by investigating the present diet in the pre-96 months group. It is possible, however, that the observed variation in physique applies only to the particular children involved at the present time. To determine which of these two alternatives is more likely would require waiting until younger children reached the 96-143 months age group. From the standpoint of time involved, that solution would be analogous to a long-term follow-up study.

Consequently, it is felt that a compromise must be made by studying the diets of those children who differ in physique now, viz., those in the 96-143 months group. To attempt an investigation into the nutritional history of this group is obviously impractical. Recognizing, however, that dietary patterns are habit-

It is also an age range in which sexual maturity is a very important factor determining growth. Indeed some or all of the between group differences in body measurements reported in this paper *might* arise from variation in the timing of the adolescent growth spurt in the

TABLE 5

Group mean bi-iliac diameter adjusted for age and height

AGE GROUPINGS	BI-ILIAC DIAMETER (CMS.) ADJUSTED FOR AGE AND HEIGHT[1]				
	A	B	A-B	F-ratio of difference	Coefficient of variation
Boys					
72-107 mos. (6 thru 8 yrs.)	20.5	20.3	0.2	1.29	4.3
96-143 mos. (8 thru 11 yrs.)	22.4	22.2	0.2	0.65	6.0
132-167 mos. (11 thru 13 yrs.)	24.1	24.5	−0.4	1.74	6.5
108-143 mos. (9 thru 11 yrs.)	22.9	22.7	0.2	0.20	6.3
Girls					
72-107 mos. (6 thru 8 yrs.)	20.1	19.8	0.3	1.67	4.8
96-143 mos. (8 thru 11 yrs.)	22.4	22.3	0.1	0.20	5.2
132-167 mos. (11 thru 13 yrs.)	24.5	24.7	−0.2	0.36	6.0
108-143 mos. (9 thru 11 yrs.)	22.7	22.7	0	0.01	5.4

[1] The means were adjusted for age and height differences indicated in Table 1.

groups compared. We are now engaged in the study of methods suitable for use in determining the mean age of the pubertal growth spurt of groups of children.

c. Indices of body measurements

Using the age classification 96-143 months, group mean heights and weights adjusted for age readily indicated differences in the growth and physique of the two groups of children. For both boys and girls, these two measurements were highly significant in the power to dis-

forming and remain relatively unchanged over short periods of time, it seems reasonable to assume that any dietary differences of former years between these two socio-economic groups would still be detectable with a fair degree of accuracy. In future work we will investigate, therefore, the diets of the 96-143 months group.

criminate between the two socio-economic groups. They are fairly simple routine measurements requiring little specialized training. From Table 2 it can be seen that children of both sexes in school A were 3 to 3½ centimeters taller, and over 8 pounds heavier than children of the same age in school B.

The other measurements—chest, bi-iliac, and calf circumference—were also significantly different between the two schools when adjusted for age. The bi-iliac differed by approximately ¾ centimeter, the chest by ½ centimeter, and the calf by 1 centimeter. These differences con-

TABLE 6

Group mean transverse chest diameter adjusted for age and height

AGE GROUPINGS	CHEST DIAMETER (cms.) ADJUSTED FOR AGE AND HEIGHT[1]				
	A	B	A-B	F-ratio of difference	Coefficient of variation
Boys					
72-107 mos. (6 thru 8 yrs.)	19.8	19.9	−0.1	0.14	5.3
96-143 mos. (8 thru 11 yrs.)	21.3	21.3	0	0.04	5.8
132-167 mos. (11 thru 13 yrs.)	22.8	23.3	−0.5	2.51	6.7
108-143 mos. (9 thru 11 yrs.)	21.6	21.5	0.1	0.18	5.6
Girls					
72-107 mos. (6 thru 8 yrs.)	19.1	19.1	0	0.01	5.6
96-143 mos. (8 thru 11 yrs.)	20.7	20.9	−0.2	0.77	7.2
132-167 mos. (11 thru 13 yrs.)	22.0	22.4	−0.4	1.58	7.8
108-143 mos. (9 thru 11 yrs.)	20.9	21.1	−0.2	0.83	7.6

[1] The means were adjusted for age and height differences indicated in Table 1.

firmed the findings on weight and height but provided little additional information of discriminating value between the two schools.

When weight was adjusted for both age and height, it was indicated in Table 3 that there no longer appeared a significant difference between the two schools. (Girls 72-107 months were an exception to this. The probable origin of this weight difference will be discussed presently.) Similarly, the chest and bi-iliac diameters were not significantly different between the two schools when adjusted for age and height. These results are given in Tables 5 and 6.

The calf circumference, nevertheless, did indicate a significant ($P = 0.02$) difference of almost 1 centimeter between boys in the two schools when adjustments were made for both age and height. The results are presented in Table 7. In fact, the difference between the

TABLE 7

Group mean calf circumference adjusted for age and height

AGE GROUPINGS	CALF CIRCUMFERENCE (cms.) ADJUSTED FOR AGE AND HEIGHT[1]				
	A	B	A-B	F-ratio of difference	Coefficient of variation
Boys					
72-107 mos. (6 thru 8 yrs.)	24.9	24.3	0.6	3.68	6.3
96-143 mos. (8 thru 11 yrs.)	27.6	26.7	0.9	5.12*	8.6
132-167 mos. (11 thru 13 yrs.)	29.8	29.6	0.2	0.19	10.0
108-143 mos. (9 thru 11 yrs.)	28.1	27.1	1.0	3.35	8.8
Girls					
72-107 mos. (6 thru 8 yrs.)	25.2	24.3	0.9	6.71*	6.0
96-143 mos. (8 thru 11 yrs.)	28.0	27.4	0.6	2.08	8.8
132-167 mos. (11 thru 13 yrs.)	30.1	30.1	0	0.01	9.3
108-143 mos. (9 thru 11 yrs.)	28.3	27.9	0.4	0.60	9.2

[1] The means were adjusted for age and height differences indicated in Table 1.
* Significant at 5% level.

mean calf circumference of the two groups of boys was reduced only slightly when the correction for height was introduced in addition to age. The difference in the girls' calf circumference between school A and school B was not as great as that of the boys in the age group 96-143 months, nor was it significant.

For girls 72-107 months, however, there was a highly significant difference of about 1 centimeter between the mean calf circumference of the two schools. It was felt that this might explain the source of the weight differential observed at this age when adjustments were made for age and height. When the two groups of girls were compared in

mean weight after adjustment for age, height, and calf, there appeared no difference whatsoever.

These results with the circumference of the calf suggest that such measurement is possibly another index useful in pointing out physical differences between socio-economic groups when adjustment is made for both age and height. Variations in the muscle mass were not always reflected in weight changes after adjustment had been made for age and height. For example, boys 96-143 months differed significantly in the mean calf circumference but not in weight when adjusted for age and height.

d. Use of factor analysis to measure an index of physical structure [3]

It was indicated in the previous section that a suitable description of the physical characteristics of a group might be obtained by limiting measurements to weight, height, and calf circumference, with corresponding adjustment for age. It is of interest to see if this recommendation can be confirmed in studying the problem from another approach.

The technique of multiple factor analysis devised by Thurstone (16) is familiar to workers in psychology and allied fields. Its applicability to a medical problem has been indicated by Andrews (1). It was probable that the method would also be of value in this instance.[4]

The purpose was to derive a composite index for measuring general physique from several variables describing certain elements of physical structure. Our desire was to determine, at least, which variables in that index are the most important or have the greatest influence.

According to the principles of factor analysis, suppose a set of n variables is associated with each individual of a population. These variables might be scores made by persons in tests of speed and ability in solving certain types of problems, or they might be various physical properties, as in the present instance. These variables are usually

[3] This investigation was supported (in part) by a research grant from the National Institutes of Health, U. S. Public Health Service.

[4] Since writing this section, the authors have received an abstract of doctoral dissertation T883, by Helen Heath, University of Chicago, on "A factor analysis of women's measurements taken for garment and pattern construction." The dissertation and references cited therein used methods of factor analysis on extensive lists of physical measurements. As far as these references could be checked, their problem was the classification of body types, an objective somewhat related to but different from the one in the present instance.

correlated among themselves to a considerable extent. The question arises whether it would be possible to reduce this set of n variables to a more fundamental and basic set, fewer in number than n, which will automatically determine the values of the n variables. There are always n possible sets, called principal components.

In psychology and educational testing, the first set or first principal component has been called a mental factor, such as general ability. It appeared reasonable to expect that by using factor analysis methods with anthropometric data that the first principal component would represent an index of general physique. If so, the first factor loadings (see a_i below) would indicate the comparative importance of each anthropometric variable in that index or component.

The philosophy behind the application of the method in this example may be summarized as follows:

Let

$$S_1 = a_1z_1 + a_2z_2 + a_3z_3 + a_4z_4 + a_5z_5$$

where

S_1 = the value obtained as a score from the first principal component, and purported to be an index or score of physical structure,

z_1 = an individual's weight expressed as a standard deviate from the group mean weight,

z_2 = an individual's height expressed as a standard deviate from the group mean height,

z_3 = an individual's bi-iliac expressed as a standard deviate from the group mean bi-iliac,

z_4 = an individual's chest expressed as a standard deviate from the group mean chest,

z_5 = an individual's calf expressed as a standard deviate from the group mean calf,

a_1 = first factor loading for weight,
a_2 = " " " " height,
a_3 = " " " " bi-iliac,
a_4 = " " " " chest, and
a_5 = " " " " calf.

As a result of expressing the n variables as standard deviates with mean zero and unit variance, the total variance of all possible principal

components can be written as n. The percentage of total variation explained by the first principal component is obviously

$$100V(S_1)/n = 100\Sigma a_i^2/n.$$

The criteria that would be used in discriminant function analysis to estimate values of the a's are not applicable in factor analysis. A different type of restriction was resorted to in this case, since no dependent variate was available.

It was assumed that a homogeneous group of individuals would have a common type of physical build, expressed by an identical score, S_1. Based upon this assumption, estimates of the a's were made so that the variation among these scores was a maximum. This was the criterion used by Hotelling (13) and Horst (12) in evaluating the coefficients in the index. The same results would be obtained, however, by using the principle of Edgerton and Kolbe (8) of minimizing the variation in score for each individual. In other words, children with similar scores were members of the same group, and the a's were chosen so that the scores were similar within a known homogeneous group.[5]

Insofar as the problem here was concerned, interest was centered on the relative magnitude of the a's for the first factor. Our primary concern was with the contribution of each of the five variables in determining the physical score.

The method of Hotelling, based on product-moment correlations of all possible combinations of the five variables, was used. The only change here was that the correlations in the matrix were partial correlation coefficients. This removed the variations caused by age in all of the zero-order correlations. This had no effect on the results or interpretation of the procedure.

The method of analysis was applied to both schools, boys and girls, and two age groups. A listing of the zero-order correlation coefficients for the five variables plus age in months is given for the age group 96-143 months in Table 8. The result of the calculations to estimate the a's is given in Table 9. The relative position of the three largest

[5] There is difficulty here in that we do not actually know how homogeneous our groups are with regard to physical structure and other characteristics. Further studies in school A have led us to believe that the group is not entirely homogeneous, although as a group the children are markedly dissimilar from those in school B.

values of the a's in each group is indicated in parentheses below the value itself.

It can be seen at a glance that a_1, the factor loading for weight, was uniformly the largest coefficient in each of the eight indexes. That is, weight of an individual was the most important measurement to describe general physique. It can be seen from the last column in

TABLE 8

Correlations among physical measurements and age, using data for schools A and B, boys and girls, aged 96-143 months

VARIABLE	AGE	WEIGHT	HEIGHT	BI-ILIAC	CHEST	CALF
Age	.....	0.5671	0.7086	0.5188	0.5430	0.5247
	.....	0.5139	0.7058	0.6285	0.4531	0.4375
Weight	0.6028	.....	0.7772	0.9096	0.8727	0.9547
	0.5655	.....	0.7064	0.8726	0.8904	0.9389
Height	0.6814	0.8359	.....	0.7702	0.7298	0.7030
	0.7583	0.7846	.....	0.8035	0.5859	0.6053
Bi-iliac	0.5747	0.7881	0.8115	.....	0.7984	0.8293
	0.6234	0.9660	0.8573	.....	0.8049	0.8040
Chest	0.4724	0.7789	0.6361	0.6250	.....	0.8432
	0.3644	0.8238	0.5635	0.7615	.....	0.8223
Calf	0.4649	0.9081	0.6260	0.6351	0.6600	.....
	0.5488	0.9512	0.7078	0.8863	0.7587	.....

Note: Upper figure in each pair is for boys.
Lower figure in each pair is for girls.
Figures above the diagonal represent school A.
Figures below the diagonal represent school B.

Table 9 that between 65 per cent and 80 per cent of the total variation in these five measurements was explained by the factor of general physical structure. Weight by itself, however, explained one-fourth of the variation found in the first principal component. That is,

$$a_1^2/\Sigma a_i^2 = 1/4,$$

on the average.

Another striking point in Table 9 was that the measurement of height, especially for school A, was relatively unimportant. Instead, the calf and bi-iliac seemed most sensitive or descriptive in that order. Among the indexes created for school B, it can be seen that height and calf were the two heavily weighted measurements following that of weight.

When the eight indexes were examined from the standpoint of sex

and age, there were two conclusions which seemed of interest. The first is that older girls, 96-143 months, appeared to be best described by weight, bi-iliac, and calf, in that order. The second was that for all boys the weight and calf were best, and the third depended on socio-economic class—for school A it was bi-iliac, and for school B it was height.

TABLE 9

First factor loadings for five physical measurements, using factor analysis on selected groups of school children

SEX	AGE (months)	FIRST FACTOR LOADING					PERCENTAGE OF TOTAL VARIATION EXPLAINED BY FIRST FACTOR
		a_1	a_2	a_3	a_4	a_5	
		Weight	Height	Bi-iliac	Chest	Calf	
				School A			
Boys	72-107	0.9615 (1)	0.8571	0.9181 (2)	0.8642	0.8876 (3)	80.7
	96-143	0.9683 (1)	0.7608	0.9092 (3)	0.8804	0.9164 (2)	79.2
Girls	72-107	0.9529 (1)	0.7794	0.7448	0.8199 (3)	0.8578 (2)	69.6
	96-143	0.9648 (1)	0.6823	0.9126 (2)	0.8839	0.9114 (3)	76.8
				School B			
Boys	72-107	0.9723 (1)	0.9066 (2)	0.8441	0.7997	0.8703 (3)	77.5
	96-143	0.9613 (1)	0.8106 (3)	0.8036	0.7700	0.8303 (2)	70.2
Girls	72-107	0.9145 (1)	0.8362 (2)	0.8199 (3)	0.7607	0.6930	65.3
	96-143	0.9816 (1)	0.7555	0.9621 (2)	0.8381	0.9096 (3)	79.8

e. Summary and discussion

Summing up, the results indicated that children 96-143 months (8 thru 11 years) of age from upper socio-economic classes were taller and weighed more for their age. They were not heavier for their age when consideration was made for the fact that they were taller. The boys in the upper socio-economic class appeared to have more muscle mass and subcutaneous fat, as judged by calf circumference, for their

age and height. This difference was not sufficiently great to cause significant weight changes. Younger girls, 72-107 months (6 thru 8 years), in the upper socio-economic group appeared to have more muscle mass and subcutaneous fat of an amount sufficiently great to cause weight differences for their age and height over those in a different socio-economic class. As a physical discriminant between socio-economic groups,[6] the chest measurement offered little additional information which was not already provided by weight and height and possibly calf circumference. The bi-iliac measurement may have additional value in certain groups, such as girls from higher socio-economic classes and aged 96-143 months (8 thru 11 years).

In future work, it is hoped to make comparisons of growth with dietary patterns and stages of puberty. As Hogben, Waterhouse, and Hogben (11) have pointed out, "Height and weight norms exhibited without regard to the level of sexual development may be deceptive as standards for growth comparisons."

REFERENCES

1. Andrews, T. Gaylord (1948). Statistical studies in allergy, Part II: A factorial analysis. *Journal of Allergy*, 19, No. 1, January, pp. 43-46.
2. Boas, F. (1928). Family traits as determined by heredity and environment. *Proceedings of the National Academy of Science*, 14, p. 496.
3. Bryan, A. H., and B. G. Greenberg (1949). Methods for studying the influence of socio-economic factors on the growth of school children. *Journal of the Elisha Mitchell Society*, 65, No. 2, December, pp. 311-314.
4. Bryan, A. H., and B. G. Greenberg (1950). Influence of social and economic factors on the growth of children. Presented before Southeastern Branch, Society for Experimental Biology and Medicine, Duke Hospital, Durham, North Carolina.
5. Bosley, Bertlyn (1947). A practical approach to nutrition education for children. *Journal of American Dietetic Association*, 23, No. 4, April, pp. 304-309.
6. Burgess, May Ayres (1937). The construction of two height charts. *Journal of American Statistical Association*, 32, No. 198, June, pp. 290-310.
7. Cox, Gertrude, Robert C. Eckhardt, and W. G. Cochran (1940). The Analysis of Lattice and Triple Lattice Experiments in Corn Varietal Tests,

[6] This use of a physical discriminant should be distinguished from the customary application of a discriminant to classify an individual into a particular group.

Part II. Mathematical Theory. *Agricultural Experiment Station Research Bulletin* 281, Iowa State College, Ames, Iowa.

8. EDGERTON, HAROLD A., and LAVERNE E. KOLBE (1936). The method of minimum variation for the combination of criteria. *Psychometrika*, 1, September, pp. 183-187.
9. FINNEY, D. J. (1946). Standard errors of yields adjusted for regression on an independent measurement. *Biometrics Bulletin*, 2, No. 3, June, pp. 53-55.
10. HARRIS, MARILYN, D. G. HORVITZ, and A. M. MOOD (1948). On the determination of sample sizes in designing experiments. *Journal of American Statistical Association*, 43, No. 243, September, pp. 391-402.
11. HOGBEN, H., J. A. H. WATERHOUSE, and L. HOGBEN (1948). Studies on puberty, Part I. *British Journal of Social Medicine*, 2, No. 1, January, pp. 29-42.
12. HORST, PAUL (1936). Obtaining a composite measure from a number of different measures of the same attribute. *Psychometrika*, 1, March, pp. 53-60.
13. HOTELLING, HAROLD (1933). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24, September-October, pp. 417-441, 498-520.
14. KEMSLEY, W. F. F. (1950). Weight and height of a population in 1943. *Annals of Eugenics*, 15, Part 2, pp. 161-183.
15. STUART, HAROLD C. (1946). Normal growth and development during adolescence. *New England Journal of Medicine*, 234, May 16, 23, 30, pp. 666-672, 693-700, 732-738.
16. THURSTONE, LOUIS L. (1947). Multiple Factor Analysis. University of Chicago Press, *Chicago*.

APPENDIX

The pilot studies upon which the sampling design was based consisted essentially of two main types. In the first series, each of ten boys, aged 9-10 years, in a nearby orphanage was subjected to physical measurement by each of two observers. The children were first dressed and then undressed. Each measurement was also done in duplicate. These measurements were used in the study of the importance of clothes for group differences as well as estimating between-observer and within-observer variabilities.

It was only at the orphanage that the children were measured without clothes. In a second series of experiments, five boy and five girls from the fourth grade of school B were submitted twice to each of two observers in random sequence for measurements of height, bi-iliac dia-

meter, transverse chest diameter and calf circumference. In the case of weight, a different observer made two independent weighings of each child. The figures obtained were studied by analysis of variance to determine estimates of the separate components of variance. It was postulated that the variation in body measurements would be a combination of the components due to individual children, variability due to different technicians, and the repeatability of a particular observation by the same technician. Estimates of the components of variance made from these studies are presented in Table 10. It should be pointed out

TABLE 10

Estimation of components of variance in measuring school children

CHARACTERISTIC AND UNIT OF MEASUREMENT	BETWEEN SUBJECTS		BETWEEN OBSERVERS		WITHIN OBSERVERS	
	Boys	Girls	Boys	Girls	Boys	Girls
Height (cms.)	59.84	63.51	0.00	0.04	0.50	0.02
Bi-iliac (cms.)	1.58	1.88	0.05	0.00	0.06	0.02
Chest (cms.)	1.03	0.52	0.01	0.28	0.04	0.05
Calf (cms.)	4.40	7.35	0.01	0.02	0.01	0.01
Weight (lbs.)	108.80	323.60	*	*	0.00	0.00

* This portion omitted from pilot study.

Note: These values were based on small numbers of degrees of freedom, and their erratic nature indicates that they should be considered only as preliminary estimates.

that these estimates are quite variable inasmuch as they are based on relatively few observations. No attempt was made to place confidence limits about these estimates.

AN APPLICATION OF STATISTICAL METHODS TO THE EXAMINATION OF SPUTUM FOR NEOPLASTIC CELLS*

BY EUGENE K. HARRIS

INTRODUCTION

EXAMINATION of various secretions for neoplastic cells has received considerable attention in the past fifteen years in view of the possibilities of diagnosis before metastasis has occurred. This is especially important in primary bronchogenic carcinoma in which only total pneumonectomy seems to offer hope of recovery.

It has been recommended that routine sputum examinations be performed on all patients suspected of carcinoma of the lung. Thorough examination of sputum smears is a time-consuming task and cannot be routinely undertaken by a busy pathologist. Therefore, suitably trained non-medical technicians must be employed to screen smears, bringing all suspicious cell groups to the attention of the pathologist for final diagnosis.

At present there exists no information concerning the diagnostic value of a single sputum specimen or smear when handled and examined by a trained technician. Moreover, the possibility of variation in the diagnoses of several technicians independently examining the same smears has not been explored. Coupled with this "inter-individual" variation is an "intra-individual" variation in diagnosis which may arise when one technician reads the same smear on two separate occasions.

The present study has investigated these problems statistically in hopes of arriving at quantitative criteria which would enable the routine cytologic examination of sputum smears to be based on an efficient design and to yield accurate results.

* This study was carried out in the Department of Public Health, with the cooperation of the Department of Pathology of Yale University, and was supported by a grant from the National Cancer Institute, U. S. Public Health Service.

The most comprehensive studies done to date on the subject of neoplastic cells in sputum have been those of Wandall (1944), Farber *et al.* (1948), Liebow *et al.* (1948), and Woolner and McDonald (1949).

All of these workers used the wet-film technique introduced by Dudgeon and Wrigley (1935), followed either by haemotoxylin-eosin or Papanicolaou polychrome staining. The studies of Farber *et al.*, and Woolner and McDonald are noteworthy because of the very large number of patients examined, the great majority believed to have non-neoplastic lesions. However, the reports of these workers indicate a lack of discrimination between results from patients known and unknown to the observer. Liebow has pointed out that the pathologist must have no knowledge of the clinical status or even the age of the patient in order to insure a completely unbiased cytological diagnosis. Only Wandall's study seems to escape this criticism.

Results of the four studies mentioned are summarized in Table 1.

TABLE 1

Review of previous results in the examination of sputum for neoplastic cells

OBSERVER	SPECIMENS PER PATIENT	SLIDES PER SPECIMEN	PATIENTS WITH CA.	PATIENTS WITHOUT CA.	CORRECT POSITIVE DIAGNOSES	FALSE POSITIVE DIAGNOSES
Wandall	3	5-6	100	93	84/100 (84%)	6/93 (6%)
Farber *et al.*	3	3-5	66	400(?)	54/66 (82%)	?
Liebow *et al.*	2-3	1	49	59	21/49 (43%)	3/59 (5%)
Woolner and McDonald	1-2	5	180(?)	1400(?)	80% *	?

* Estimated by the authors.

MATERIALS AND METHODS

The wet-film method of preparing sputum smears was used throughout the present investigation. The patient was given a sterile Petri dish and carefully instructed to raise sputum from "deep down in the chest"—the trachea and bronchi. He was asked to rinse his mouth with water some minutes prior to coughing. This helped to remove food particles and superficial cellular detritus from the oral cavity. In the present study all collections were made in the morning between nine and eleven o'clock.

Dishes were gathered in about fifteen minutes and immediately taken to the laboratory. The sample of sputum was inspected against a black background to facilitate the separation of solid or blood-tinged portions from the surrounding mucous. With a wire loop likely bits of sputum were gently smeared in a thin film on a clean glass slide. It was necessary to flame the loop between transfers.

Smears should not be made by rubbing sputum between two slides for there is then danger of disrupting cell groups and injuring individual cells. Blood-tinged portions of the specimen should always be selected since these are most apt to harbor neoplastic cells. If the sputum is not blood-streaked, the slides should represent random selections from solid material in different regions of the dish. In this study three slides were made from each specimen.

The slides were immersed *while wet* in a solution of equal parts of ether and absolute alcohol. This fixative was first suggested by Papanicolaou (1933) for use in preparing vaginal smears. Slides should be kept in the fixative for at least one-half hour; practice in this laboratory has shown that they may be left as long as a week without harmful effects.

The staining procedure represents a slight modification of that originally introduced by Papanicolaou (1942):

1. Pass slides through successive changes of 95 per cent, 80 per cent, and 50 per cent alcohol (two minutes in each change).
2. Stain for 5½ minutes in filtered aqueous solution of Harris haemotoxylin.
3. Rinse for one minute in distilled water.
4. "Blue" for one minute in 0.1 per cent ammonia water.
5. Rinse again for one minute in distilled water.
6. Dehydrate in successive changes of 50 per cent and 80 per cent alcohol (one minute in each change).
7. Stain for 1½ minutes in unfiltered Orange G solution, consisting of:

Orange G	1.00 gms.
Phosphotungstic acid	0.05 gms.
95 per cent alcohol	200 cc.

8. Rinse for one minute in 95 per cent alcohol.
9. Stain for two minutes in unfiltered EA-50 solution consisting of:

Light green	1.00 gms.
Bismarck brown	0.16 gms.
Eosin Y	0.84 gms.
Phosphotungstic acid	0.68 gms.
95 per cent alcohol	380 cc.
Distilled water	20 cc.

10. Dehydrate in two changes of 95 per cent alcohol and two changes of absolute alcohol (one minute in each change).

11. Clear in two changes of xylol (five to fifteen minutes in xylol is suitable).

12. Mount in Permount. A 20- or 22-mm. × 40 mm. cover slip may be used.

After staining, the slides were submitted to systematic microscopic study with the help of a mechanical stage. In this way every cell field under the cover slip was scrutinized. Low power magnification (× 120) sufficed for the bulk of the observation, but all suspicious fields were subjected to higher power (× 440) before an opinion was reached. All atypical cells or cell groups which appeared as if they might come from a carcinoma were shown to the consulting pathologist for confirmation.

In this study each slide was identified by number at the time the smear was made. However, examination took place from one to several days after the recording of the number against the name of the patient. By this time the observer had lost all remembrance of number-patient correspondence. Moreover, slides were examined in batches of twenty, each batch including specimens drawn from several patients. As an added precaution, the numbers were masked with gummed labels and the slides rearranged in random order just before microscopy. Thus, the writer feels justified in claiming that, at the time of examination, he had not the slightest information about the source of any of the slides, nor could the opinion reached on one slide possibly influence his decision on another slide from the same specimen.

It has been stated that one of the objects of this investigation was to determine inter-individual and intra-individual variation among technicians trained to read sputum smears routinely. To this end, three such technicians agreed to cooperate with the writer in forming a group of four non-medical observers with special skill in sputum cytology but otherwise no formal experience in pathology.

One of these technicians had had sufficient experience with sputum smears in the pathology laboratory of the Hartford Hospital, Hartford, Connecticut, to allow admission to the study without further examination. A second technician had less experience with the technique but after some preliminary training was able to report on a test set of slides with sufficient accuracy to insure reliable data in the study proper.

One other technician, however, was completely new to the technique and passed through a training period of about three months. During this time she received intensive instruction and experience in the collection, preparation, and examination of sputum specimens and was

acquainted with the various clinical and surgical aspects of cancer of the lung. At the end of three months, following a satisfactory performance on a final test set of slides, she was admitted to the study.

Each of the three cooperating technicians has independently read and reported on one hundred slides selected for distribution. Moreover, two technicians examined each slide twice, the second examination separated from the first by an interval of approximately two months. These technicians were unaware of the fact that they were making repeat examinations of each slide but rather were under the impression that they were being asked to report on two hundred different slides. Slides were distributed in random order in batches of twenty and their identities masked by code numbers so that the reading on one slide could in no case affect a subsequent examination of a slide from the same specimen.

The writer, on previous examination, had noted positive fields, confirmed by the pathologist, on thirty-four of the one hundred slides selected. The remaining sixty-six were thought to be negative. However, since the writer felt that, acting as liaison between the other technicians, he had become too familiar with this set of slides, duplicate readings were made with a separate group of one hundred slides. In this case a period of some six months had elapsed between readings. The usual precautions of randomization and coding were taken. Any field which a technician thought positive or very suspicious was brought to the attention of the pathologist for confirmation. If the pathologist considered the field non-neoplastic, the slide as read by that technician was designated negative.

This circumstance leads to an important conclusion which should be emphasized. It is impossible for a technician to be charged with a false positive slide since final judgment on any suspicious field must come from the pathologist. Therefore, he alone is responsible for any false positive reported. On the other hand, the occurrence of a false negative slide (i. e., a slide containing neoplastic cells which has been incorrectly diagnosed negative) will usually be the fault of a technician who, in examining the slide, has missed a positive field present on it. Such a false negative is revealed only when another technician finds a positive field on the slide which has been called negative, or when the technician originally at fault discovers a positive field upon a subsequent re-examination of the same slide. Of course, on occasion a false negative will occur when a pathologist incorrectly reverses the decision of a technician who has brought to his attention a field actually positive.

In short, the function of the technical aide in sputum cytology work is to screen the great mass of negative smears and present to the pathologist only those cell groups which appear suspicious or frankly positive. Naturally, neither person is infallible. These considerations will be enlarged upon later in this paper.

The cytology of sputum has been described most exhaustively by Wandall although every recent report on the examination of sputum in the diagnosis of lung tumor has contained a review of the various types of cells encountered. Since any realistic description of these cells must of necessity be lengthy and detailed, the subject has been omitted entirely in an effort to focus attention on the statistical methods employed in evaluating sputum diagnosis. A fairly complete discussion of the cellular content of sputum, aided by microphotographs, has been included elsewhere (Harris, 1950).

THE DIAGNOSTIC VALUE OF SPUTUM EXAMINATION

The present study includes material drawn during the course of about eighteen months from 152 patients who were hospitalized on the chest service of the Grace-New Haven Community Hospital. The presence of carcinoma of the lung has been confirmed in 63 of this group by microscopic examination of tissue received at bronchoscopic biopsy, operation, or autopsy. In five other patients no tissue diagnosis of cancer was available, but clinical and roentgenographic examinations leaned so heavily in that direction that these patients have been included in the group positive for cancer.

Six patients were proven by tissue diagnosis to be suffering from cancer of the esophagus. Since these patients yielded sputum which was in every case negligible in amount and unfit for examination, they have been omitted from the study.

Metastatic carcinoma of the lung has been proven in five patients. The primary sites were, respectively, the pharynx, prostate, rectosigmoid, cervix, and liver. In one case neoplastic cells were found in the sputum. However, it cannot be assumed that sputum diagnosis failed in the other four patients since the metastatic growth may not have communicated with a bronchus. In fairness, therefore, all five patients were omitted from the group considered positive for bronchogenic carcinoma.

In forty cases the presence of chest pathology other than carcinoma was confirmed by examination of specimens removed at operation. In 27 other cases where no microscopic proof was available the previous

history of the patient, the clinical and roentgenographic findings, and the subsequent course of the disease strongly indicated that the patient was suffering from non-neoplastic disease.

Finally, there remains a small nucleus of 6 patients about whom no final conclusion can be reached. In three cases bronchoscopy revealed masses in the main or intermediate bronchi, but biopsies in two of these patients were reported showing only chronic inflammatory reaction and in the third patient merely bronchial mucosa. In all 6 cases a clinical or roentgenographic examination which admitted the possibility of tumor diminished in importance in view of a favorable hospital course and follow-up in the clinic. These individuals have been under observation for lengths of time varying from twelve to twenty months. In no case has a death certificate been received.

An effort was made to obtain three sputum specimens from each patient included in the study. In some cases this was impossible because of early discharge or because the patient underwent surgery almost immediately after admission. Three specimens were obtained from 63 of the 68 patients positive for cancer of the lung. Three patients yielded two specimens while only single specimens were obtained from the remaining two patients. The total number of sputum specimens obtained from these patients was 197.

It would be unwise to estimate statistically the results of missing specimens in the two cases in which only one specimen was obtained. This would in fact be "filling in" two-thirds of the data. However, for some of the analyses to follow in this chapter, it would be helpful to estimate the result of the missing third specimen in those cases in which no more than two specimens were obtained.

In order to estimate the probable result of a missing third specimen, the patient in question may be compared with other patients with cancer who yielded three specimens. In each of the three patients under consideration the two specimens obtained were both positive. Investigation revealed thirteen patients yielding three sputa of which the first two were likewise positive. The third specimen was reported positive in nine of these cases, negative in four. Applying a similar ratio to the three patients with missing specimens, it will be assumed that a third specimen would have been positive in two of these cases and negative in one. No attempt will be made to estimate the results of particular slides from these missing specimens.

The data which now follow concern the number of positive specimens

and slides obtained from cases of proven cancer. These figures should be understood as the results of microscopic examinations by a well-trained technician subject to the human variations analyzed later in this paper.

Table 2 shows the number of patients with cancer yielding one, two, or three positive sputum specimens. These data include the estimated results of missing third specimens. However, the two patients who yielded only single specimens are not included.

TABLE 2

Numbers of patients with carcinoma yielding 0, 1, 2, or 3 positive sputum specimens

NUMBER OF POSITIVE SPECIMENS	NUMBER OF PATIENTS	
	Observed	Calculated
0	31	18.7
1	13	29.3
2	11	15.3
3	11	2.7
Total	66	66.0

Thirty-one patients did not offer any positive specimens. Therefore, 53 per cent of a total of 66 patients with carcinoma of the lung yielded at least one positive sputum specimen.

False positive smears were obtained in four patients. In two of these cases, only one of the nine smears examined was incorrectly diagnosed. The percentage of patients without neoplastic disease of the lung who were incorrectly reported positive is $4/67 \times 100$, or 6.0 per cent.

Before attacking some of the larger problems posed by the data thus far presented, two questions may be resolved by routine Chi-square examinations. We wish to determine whether any one of three morning specimens obtained from a patient with cancer of the lung is especially apt to harbor neoplastic cells. Table 3 gives the number of positive first, second, and third specimens obtained in this study. It is assumed of course, that all specimens were collected, processed, and examined with equal care.

TABLE 3

Numbers of positive first, second, and third specimens obtained

NO. OF POSITIVE 1ST SPECIMENS	NO. OF POSITIVE 2ND SPECIMENS	NO. OF POSITIVE 3RD SPECIMENS
22	25	21
	Total: 68	

If, indeed, no one specimen is especially apt to contain neoplastic cells, then each category of Table 3 should include one-third of the total, or 22.67. The Chi-square (χ^2) test judges the significance of the discrepancy between the observed and expected situations:

$$\chi^2 = \Sigma\left[\frac{(O-E)^2}{E}\right] = \frac{(0.67)^2}{22.67} + \frac{(2.33)^2}{22.67} + \frac{(1.67)^2}{22.67}$$

$$= 0.38, \text{ with two degrees of freedom.}$$

The probability of the chance occurrence of such a value is .83. It is clear that all specimens are equally capable of containing neoplastic cells.

Exactly the same question may be asked of the three slides made from a single specimen. Table 4 exhibits the requisite data for first, second, and third slides. The same assumptions hold here as in the analysis of specimens.

TABLE 4

Numbers of positive first, second, and third slides obtained

NO. OF POSITIVE 1ST SLIDES	NO. OF POSITIVE 2ND SLIDES	NO. OF POSITIVE 3RD SLIDES
40	42	46
	Total: 128	

The expected number in each category of Table 4 is 42.67. Then,

$$\chi^2 = \frac{(2.67)^2}{42.67} + \frac{(0.67)^2}{42.67} + \frac{(3.33)^2}{42.67} = 0.44.$$

The probability attached to this value of χ^2 is .80. Clearly no one slide is more likely than another to contain tumor cells.

At this point the data of Table 2 may be examined to estimate the probability that a single specimen from a cancer patient will reveal neoplastic cells. We have a total of 198 specimens, of which 68

$(= 13.1 + 11.2 + 11.3)$ are positive. The chance that a specimen taken at random is positive may thus be estimated as

$$P = 68/198 = .343.$$

It is, however, to be noted that this estimate has tacitly assumed that the probability of a positive specimen is the same for all patients. This can be tested by comparing the observed distribution of Table 2 with that calculated from the binomial

$$(.343 + .657)^3.$$

The last column of Table 2 shows this comparison. It is at once evident that the observed distribution differs strongly from the calculated, in that there are too many observed cases of no positive specimens, and of three positive specimens. This suggests strongly that (as one might reasonably expect) there are considerable differences between patients in their chances of producing specimens containing neoplastic cells.

The next problem is to estimate the probability that any one of three smears prepared from a single positive specimen will contain neoplastic cells. Table 5 gives the frequency of specimens with one, two, or three positive smears.

TABLE 5

Numbers of specimens with one, two, or three positive smears

	NUMBER OF POSITIVE SMEARS			
	1	2	3	Total
Number of specimens	25	26	17	68

The method of maximum likelihood will be employed. However, the mathemtical model will be modified somewhat since the discovery of a positive sputum specimen requires that at least one positive smear be reported. Therefore, in the case of a *known* positive specimen the probability of at least one positive smear must be unity. It is very likely, however, that many specimens contain neoplastic cells but yield no positive smears. The adjusted model will take into account undiscovered positive specimens.

Let a_k $(k = 1, \cdots, n)$ represent the observed frequency of k positive smears out of n (here, $n = 3$). The probability that the number x of positive smears equals a given number k is expressed by the binomial formula:

$$P[x = k] = \frac{n!}{k!(n-k)!}[p^k(1-p)^{n-k}]. \tag{1}$$

Now,

$$P[x = 0] = (1-p)^n,$$

and

$$P[x \neq 0] = 1-(1-p)^n. \tag{2}$$

Then, the probability that $x = k$, given that $x \neq 0$ (i. e., the situation in this problem), may be expressed as

$$P[x = k \mid x \neq 0] = \frac{C_n^k[p^k(1-p)^{n-k}]}{1-(1-p)^n} .^1 \tag{3}$$

The probability of obtaining the observed distribution of frequencies is given by the expression

$$\mathbf{P}[E] = C_1 \prod_k \left[\frac{p^k(1-p)^{n-k}}{1-(1-p)^n}\right]^{a_k}. \tag{4}$$

Differentiating logarithmically with respect to the parameter p and setting the result equal to zero, we get, after some rearrangement,

$$\frac{(1-p)\Sigma ka_k - p\Sigma(n-k)a_k}{p} = \frac{n(1-p)^n \Sigma a_k}{1-(1-p)^n}. \tag{5}$$

Substituting the values, a_k, from Table 5, for $n = 3$, and transposing, one obtains

$$\frac{128(1-p)-76p}{p} - \frac{204(1-p)^3}{1-(1-p)^3} = 0. \tag{6}$$

Estimates of p may be obtained to any desired precision by a process of iteration. Trial values of p are substituted into (6) and the left-hand side calculated. The resulting value determines the next estimate of p to be interpolated. The details are given in Table 6.

TABLE 6

The iteration solution of equation 6

p	.5	.6	.58	.582	.5815
$\frac{128(1-p)-76p}{p}$	52.00	9.33	16.69	15.931	16.1204
$\frac{204(1-p)^3}{1-(1-p)^3}$	29.14	13.95	16.32	16.072	16.1360
Difference	22.86	— 4.62	0.37	— 0.141	— 0.0156

[1] This formula is a simplified expression of the well-known "Bayes' formula for probabilities of hypotheses," a discussion of which may be found in Uspensky's textbook on mathematical probability (1937).

Correct to three decimal places,

$$p = .581.$$

This is an estimate of the probability that neoplastic cells will be transferred from a specimen to any one smear made from a small portion of that specimen.

SOME PREVIOUS STUDIES OF HUMAN VARIATION

The experimental methods employed in the present study of human variation were derived in large part from two recent papers, one by Yerushalmy (1947), the other by Birkelo, Yerushalmy, *et al.* (1947).[2]

Yerushalmy and his associates discussed subjective variations in medical diagnosis with special reference to roentgenographic techniques used in tuberculosis case-finding surveys. The material for the study was drawn from a group of 1,256 individuals who were X-rayed consecutively on four different X-ray techniques, yielding for each person a 35 mm. photofluorogram of the chest, a 4″ × 10″ stereophotofluorogram, a roentgenogram on 14″ × 17″ paper negative, and a conventional 14″ × 17″ celluloid film.

These four sets of chest films of different sizes were distributed among five expert radiologists who independently diagnosed each film. Analysis of the results afforded a measure of " inter-individual " error, or " the inconsistency of interpretation found among different readers." In a typical case, one reader called 59 films positive for tuberculosis. Of these, the other observers read, respectively, 31 per cent, 29 per cent, 37 per cent, and 14 per cent negative for tuberculosis, although in many of these cases the presence of some non-tuberculous lesion was indicated (e. g., pulmonary scar, cancer, etc.). In an average of 12 per cent of the films the diagnoses varied from positive for tuberculosis to completely negative.

Following an interval of two to three months, the 14″ × 17″ celluloid films were re-circulated among the radiologists and results obtained which measure the " intra-individual " error, or the " failure of a reader to be consistent with himself in two independent interpretations of the same set of films." In the case of a radiologist who struck about the average in consistency, the writers showed that of 83 films which this

[2] For a review of earlier studies of subjective variations in diagnosis, see Harris, 1950, Ch. VI.

radiologist called positive for tuberculosis on the first reading, 18 per cent were diagnosed negative for tuberculosis on the second reading, two-thirds of these completely negative. On the other hand, this observer diagnosed 88 films positive for tuberculosis on the second reading and of these 29 per cent had been called negative for tuberculosis on the first reading.

After discussing several methods of analysis, Yerushalmy and his associates concluded that no one technique was more efficient than another in detecting cases of tuberculosis. They recommended that in mass survey work all films be read independently by at least two interpreters.

THE ACCURACY OF CYTOLOGIC TECHNICIANS

The present analysis was made to estimate the probability that a non-medical technician well-trained in cytologic diagnosis will detect neoplastic cells contained in a sputum smear. As mentioned previously, four such persons independently examined a group of one hundred sputum smears. Each smear suspected of containing neoplastic cells was brought to the consulting pathologist for confirmation.

Table 7 reflects the relatively large variation exhibited by the four technicians, A, B, C, and D. The second column of this table gives the

TABLE 7

Variation shown by technicians in diagnosing one hundred sputum smears

TECHNICIAN	NUMBER OF POSITIVE SMEARS	PERCENTAGE READ POSITIVE BY A	B	C	D
A	34		62	53	29
B	32	66		69	22
C	33	55	67		30
D	16	62	44	62	

number of confirmed positive smears detected by each observer on the first reading. The remainder of the table lists the *percentage* of those smears called positive by each of the other independent observers. Thus, for example, Technician B diagnosed 32 slides positive. Of these, Technician A called 66 per cent positive, Technician C called 69 per cent positive, and Technician D called 22 per cent positive.

The mathematical model used to analyze differences between technicians in the interpretation of sputum smears was introduced by

Neyman (1947) in a treatment of the roentgenographic data presented by Yerushalmy and his associates.

Consider a population of sputum smears divided into two exclusive categories: (1) smears showing no evidence of neoplastic growth, whose proportion is α; (2) smears positive for neoplastic cells, whose proportion is $\beta = 1 - \alpha$.

The parameters π and p are postulated to be the probabilities of correct diagnosis for categories (1) and (2), respectively. Hence, $1 - \pi$ represents the probability of a non-neoplastic smear being diagnosed as containing neoplastic cells (i. e., the probability of a false positive). Similarly, $1 - p$ denotes the probability of missing a smear containing neoplastic cells (i. e., the probability of a false negative).

Applying the notation used by Neyman, let N represent the number of sputum smears included in the sample and n the number of independent diagnoses made for each smear. Denote by k the number of positive diagnoses made for each smear. Then k may assume any one of the $n + 1$ values, $0, 1, 2, \cdots, n$. Let N_k stand for the number of smears in the sample for which there were exactly k positive diagnoses. The relative frequencies, q_k, of exactly k positive diagnoses are obtained by dividing N_k/N. There are $n + 1$ such frequencies, but only n of them are independent since they must add to unity.

Table 8 gives the observed frequencies of smears with k positive diagnoses.

TABLE 8

Absolute and relative observed frequencies of sputum smears with k $(0, 1, \cdots, 4)$ positive diagnoses

k	ABSOLUTE FREQUENCY	RELATIVE FREQUENCY
0	46	.46
1	17	.17
2	16	.16
3	18	.18
4	3	.03
Total	100	1.00

In the construction of a mathematical model, three independent parameters, α, π, and p, have been postulated. Moreover, the goodness with which the model fits the observed frequencies must be tested. Therefore, at least four independent diagnoses are required for each smear. The results obtained from the four technicians cooperating in

this study satisfy this requirement assuming that all technicians are equally competent to report on sputum smears.

Then, the probability that n diagnoses will give exactly k positive results is given by

$$P_k = \frac{n!}{k!(n-k)!}[\alpha(1-\pi)^k\pi^{n-k} + (1-\alpha)p^k(1-p)^{n-k}, \text{ for } k < n,$$

and

$$P_n = 1 - \alpha[1-(1-\pi)^n] - (1-\alpha)[1-p^n]. \tag{7}$$

Using these $n+1$ equations, the three unknown parameters may be estimated according to the method of best asymptotically normal estimates described by Neyman (1945-46).

By eliminating the three unknown parameters from the system of four independent equations (7) an equation may be obtained of the form

$$F_t(P_k) = 0, \tag{8}$$

which serves as a final restriction on the parameters.

For $n = 4$, this equation must be

$$72P_0P_2P_4 + 9P_1P_2P_3 - 2P_2^3 - 27P_0P_3^2 - 27P_1^2P_4 = 0.\ ^3 \tag{9}$$

If the right-hand sides of the system of equations (7) are substituted for the probabilities, P_i, in (9), it may be shown that (9) represents an identity in the parameters α, π, and p.

Using (9), one immediately obtains

$$F_t(q_i) = 72q_0q_2q_4 + 9q_1q_2q_3 - 2q_2^3 - 27q_0q_3^2 - 27q_1^2q_4 \tag{10}$$

in terms of the observed relative frequencies given in Table 8.

The discrepancy between the "observed" equation, $F_t(q_i)$, and the "expected" model, $F_t(P_k)$, may be measured by expanding the latter in a Taylor series about the point $P_k = q_i$. Thus,

$$F_t(P_k) = F_t(q_i) + \sum_{i=1}^{n} b_{t,i}(P_k - q_i) + R.$$

The remainder R tends in probability to zero as $N \to \infty$. The coefficients, $b_{t,i}$, represent the first order partial derivatives of $F_t(P_k)$ with respect to the expected frequencies, P_k. Neyman (1945-46) has

[3] The writer is indebted to Dr. Werner Leimbacher of the Statistical Laboratory of the University of California for the mathematical derivation of the equation (9). This derivation is reproduced in the Appendix.

shown, however, that these partial derivatives have the same values as the first order partial derivatives, $a_{t,i}$, obtainable from the expression $F_t(q_i)$ at the point $P_k = q_i$.

The terms, $a_{t,i}(P_k - q_i)$, may be obtained by differentiating the expression $F_t(q_i)$ with respect to the observed frequencies, q_i. In keeping with the hypothesis $P_k = q_i$, one may assume that

$$\bar{a} = \sum_{i=1}^{n} q_i a_{t,i} = 0.$$

Then, the expression

$$N\left[\frac{F_t(q_i)^2}{{\sigma_a}^2}\right] \tag{11}$$

has a Chi-square (χ^2) distribution with one degree of freedom. This distribution is used to test the hypothesis that the system of equations (7) with three unknown parameters adequately explains the observed frequencies.

Inserting the numerical values from Table 8,

$$\begin{aligned} F_t(q_i) &= 0.158976 + 0.044064 - 0.008192 - 0.402408 - 0.023409 \\ &= -0.230969. \end{aligned}$$

The partial derivatives, a_i, $(i = 0, 1, \cdots, 4)$, are then computed:

$$\begin{aligned} a_0 &= 72q_2q_4 - 27{q_3}^2 = -0.5292, \\ a_1 &= 9q_2q_3 - 54q_1q_4 = -0.0162, \\ a_2 &= 72q_0q_4 + 9q_1q_3 - 6{q_2}^2 = 1.1154, \\ a_3 &= 9q_1q_2 - 54q_0q_3 = -4.2264, \\ a_4 &= 72q_0q_2 - 27{q_1}^2 = 4.5189. \end{aligned}$$

Hence,

$$\bar{a} = \Sigma\, q_i a_i = -0.449475,$$

and

$${\sigma_a}^2 = \Sigma\, q_i {a_i}^2 - \bar{a}^2 = 3.824931.$$

Then,

$$\chi^2 = N\left[\frac{F_t(q_i)^2}{{\sigma_a}^2}\right] = 100\left[\frac{0.053347}{3.824931}\right] = 1.395,$$

with one degree of freedom.

The probability of the chance occurrence of a χ^2 value greater than this is .24. Evidently the proposed model (7) adequately represents the mechanism underlying the observed frequencies.

In order to estimate the parameters α, π, and p, approximations $\tilde{P}_i$ of the P_k in equations (7) are computed:

$$\tilde{P}_i = q_i[1 - \frac{a_i - \bar{a}}{{\sigma_a}^2} \cdot F_i(q_i)], \qquad i = 0, 1, \cdots, 4. \tag{12}$$

Substituting numerical values:

$$\tilde{P}_0 = .46(1 - 0.004814) = .457786,$$
$$\tilde{P}_1 = .17(1 + 0.026163) = .174448,$$
$$\tilde{P}_2 = .16(1 + 0.094495) = .175119,$$
$$\tilde{P}_3 = .18(1 - 0.228070) = .138947,$$
$$\tilde{P}_4 = .03(1 + 0.300016) = .039000.$$

In contrast to the observed frequencies, q_i, these estimated frequencies, $\tilde{P}_i$, are independent and therefore do not have to add to unity.

The case $n = 4$, requires computing the coefficients:

$$A = \frac{2(6P_0P_3 - P_1P_2)}{8P_0P_2 - 3{P_1}^2} = 1.276631,$$

$$B = \frac{9P_1P_3 - 4{P_2}^2}{3(8P_0P_2 - 3{P_1}^2)} = 0.057865,$$

and solving the quadratic

$$x^2 - Ax + B = 0,$$

with x_1 the smaller solution and x_2 the larger solution.

These solutions are

$$x_1 = 0.047061,$$
$$x_2 = 1.229570.$$

Then,

$$\hat{\pi} = \frac{1}{1 + x_1} = .955054 \qquad (= .955),$$

and

$$\hat{p} = \frac{x_2}{1 + x_2} = .551483 \qquad (= .551).$$

The parameter $\hat{\alpha}$ is obtained from the formula

$$\hat{\alpha} = \frac{4P_0x_2 - P_1}{4(x_2 - x_1)}(1 + x_1)^4 = .527807 \qquad (= .528).$$

The estimate obtained for π indicates that a proportion of 4.5 per

cent false positive diagnoses may be expected when suspicious smears are brought to a consulting pathologist for confirmation. This figure agrees well with percentages obtained in previous investigations in which the pathologist diagnosed smears without preliminary screening by technicians. As mentioned previously, the pathologist alone is responsible for false positive diagnoses.

The estimate

$$p = .551$$

means that if any one of the four technicians in this study were given one hundred sputum smears each showing some evidence of neoplasia. about 55 positive diagnoses might be expected.

STANDARD ERRORS OF THE PARAMETERS

Let $\hat{\theta}_1(q)$, a function of $q_0, q_1, \cdots, q_4$, be a BAN estimate of θ_1. Then, Taylor's formula gives

$$\hat{\theta}_1 = \theta_1 + \sum_{k=0}^{4} a_k(q_k - p_k) + R,$$

where

$$a_k = \left.\frac{\partial \hat{\theta}_1}{\partial q_k}\right|_{q_k = p_k}$$

and where the remainder R tends in probability to zero as $N \to \infty$.

Then, the asymptotic variance of $\hat{\theta}_1$ is given by, say,

$$\sigma_e^2 = \frac{1}{N} \sum_{k=0}^{4} a_k^2 p_k.$$

However, the computation of the partial derivatives, a_k, is often quite laborious, and Neyman has indicated a much more convenient method of obtaining the variance.

As illustrated by equations (7), P_k is a function of three independent parameters θ_1, θ_2, θ_3,

$$P_k = f_k(\theta_1, \theta_2, \theta_3).$$

Let

$$\frac{\partial f_k}{\partial \theta_l} = f_{k,l},$$

and

$$G_{l,m} = \sum_{k=0}^{4} \frac{f_{k,l} f_{k,m}}{f_k} = G_{m,l}, \text{ for } l, m = 1, 2, 3.$$

Let

$$\Delta = \begin{vmatrix} G_{11} & G_{12} & G_{13} \\ G_{21} & G_{22} & G_{23} \\ G_{31} & G_{32} & G_{33} \end{vmatrix}$$

and denote by $\Delta_{l,m}$ the co-factor of Δ corresponding to $G_{l,m}$.

Then, Neyman has shown that

$$\left.\frac{\partial \hat{\theta}_1}{\partial q_k}\right|_{q_k = p_k} = a_k = \frac{\sum_{l=1}^{3} f_{k,l}\Delta_{1,l}}{f_k \Delta},$$

and the asymptotic variance of $\hat{\theta}_1$ is given by

$$\sigma_e^2 = \frac{\Delta_{1,1}}{N\Delta}.$$

Let $\hat{\theta}_1 = \hat{\pi}$, $\hat{\theta}_2 = \hat{p}$, $\hat{\theta}_3 = \hat{\alpha}$.

Then, calculating the partial derivatives of P_k with respect to π, p, and α, and substituting $\hat{\pi} = \pi$, $\hat{p} = p$, $\hat{\alpha} = \alpha$ and $\hat{P}_k = f_k$, one obtains

$$\begin{aligned} \Delta &= 120.1825, \\ \Delta_{1,1} &= 13.5681, \\ \Delta_{2,2} &= 48.7523, \\ \Delta_{3,3} &= 107.1123, \end{aligned}$$

and, for $N = 100$, $\sigma_{\hat{\pi}} = \pm .034$, $\sigma_{\hat{p}} = \pm .064$, $\sigma_{\hat{\alpha}} = \pm .094$.

INTRA-INDIVIDUAL VARIATION

As described above, three of the four observers, including the writer, re-examined selected sets of one hundred smears which had been read for the first time about three months previous. In a study of variation "within" observers, or the consistency of an observer's diagnosis, one is especially interested in reversals of diagnosis from positive to negative and vice versa. Table 9 shows the frequency with which these reversals occurred.

Table 9 indicates that two observers diagnosed more slides positive on the second reading than on the first reading, while in the case of one observer the reverse situation obtained. Therefore, it cannot be

maintained that during the course of the study the observers showed significant improvement in ability to detect smears containing neoplastic cells.

TABLE 9

Reversals of diagnoses as a measure of intra-individual variation

OBSERVER	NO. OF REVERSED DIAGNOSES		TOTAL	PERCENTAGE REVERSAL (100 smears)
	Pos. → Neg.	Neg. → Pos.		
A	5	9	14	14%
B	8	19	27	27%
C	12	9	21	21%

Based on a small number of observers, a mean percentage reversal of 21 per cent is indicated with a standard deviation of $\pm$ 6.5 per cent. Considering the estimated value, p, the mean percentage reversal is not surprising. It is natural that on each successive reading a technician will discover positive smears missed at previous examinations and, at the same time, fail to report smears previously called positive. This measure of intra-individual variation in no way invalidates the estimates of the parameters α, π, and p.

DISCUSSION AND RECOMMENDATIONS

The aim of this section is to recommend some applications of the statistical results to the routine examination of sputum for neoplastic cells. Material for these examinations will be drawn from two classes of subjects, (1) hospitalized patients suspected of some chest pathology, possibly carcinoma, and (2) ambulatory persons appearing at a clinic for reasons perhaps unrelated to bronchogenic carcinoma but having a productive cough which gains the attention of the physician.

Sputum examined in the present study was collected entirely from hospitalized patients. In all the cases in which neoplastic cells were found in the sputum, clinical and roentgenographic findings had already supported a suspicion of carcinoma. In only 17 per cent of these cases was the lesion operable.

The greatest value of sputum diagnosis lies in the microscopic proof of early cancer. Therefore, for the most gratifying results, sputum examinations for neoplastic cells should be performed on every clinic patient with a productive cough. The results of the present study cannot be applied indiscriminately to ambulatory persons, and further research

is necessary before sputum diagnosis may be evaluated with regard to this group. Certainly in such a program the ratio of negative to positive smears would be overwhelming and the time and efforts of cytologic technicians would not be spent economically.

A valuable undertaking would be an investigation into the accuracy of pathologists in judging sputum smears. A consulting pathologist must bear the blame for some of the false negative as well as all of the false positive diagnoses. The present study may serve as a model for an assessment of sputum diagnoses made by pathologists.

An analysis given in the previous section indicated that one technician will *fail* to detect positive smears about 45 per cent of the time. Further calculation shows that 69 per cent of the positive smears will be correctly reported when two technicians independently read the same slide. Sensitivity (i. e., the probability of detecting positive smears) rises to 86 per cent when three technicians report independently on one slide and to 94 per cent when four independent diagnoses are made.

Since the examination of a sputum smear is a very time-consuming task, no more than two technicians can be expected to examine each smear. The increase in sensitivity obtained when two technicians work independently suffices to raise the probabilities calculated earlier in this paper. The probability that one of three specimens from a patient with carcinoma will be positive is raised from .343 to .422. Similarly, the probability that one of three smears from a positive specimen will be positive is increased from .581 to .602.

In view of these adjusted probabilities, applicable to hospitalized patients, it seems advantageous to collect at least four specimens from each patient and prepare two slides from each specimen. The time spent in collecting specimens is almost negligible. On the other hand, twenty to thirty minutes are required to examine systematically each slide. Therefore, when sputum diagnosis is contemplated as a routine procedure, the number of slides to be examined assumes great importance. Examination of two slides from each of four specimens will require less time than three slides from each of three specimens and will probably uncover more cases of bronchogenic carcinoma.

It is further recommended that two technicians independently report on each smear prepared. In order to assume a volume of work sufficient for a group of technicians, the cytologic laboratory should be situated in a large hospital which can draw material from private physicians and smaller institutions within a wide radius. Lastly, a pathologist with

special training in sputum cytology must be present to render final diagnosis.

ACKNOWLEDGEMENTS

The writer would like to express his gratitude to Dr. Averill A. Liebow, Associte Professor of Pathology, Yale University School of Medicine, who served as research advisor in pathology during the course of this study.

To Professor Jerzy Neyman of the Statistical Laboratory, University of California, and to Dr. David F. Votaw, Assistant Professor of Mathematics, Yale University, the writer is indebted for keen interest and assistance in the mathematical aspects of the work.

REFERENCES CITED

BIRKELO, C. C., E. W. CHAMBERLAIN, P. S. PHELPS, P. E. SCHOOLS, D. ZACKS, and J. YERUSHALMY. Tuberculosis case finding: A comparison of the effectiveness of various roentgenographic and photofluorographic methods. *J. A. M. A.*, 133: 359-366, Feb. 8, 1947.

DUDGEON, L. S. and C. H. WRIGLEY. On the demonstration of particles of malignant growth in the sputum by means of the wet film method. *J. Laryng. and Otol.*, 50: 752-763, 1935.

FARBER, S. M., M. A. BENIOFF, J. K. FROST, M. ROSENTHAL and G. TOBIAS. Cytologic studies of sputum and bronchial secretions in primary carcinoma of the lung. *Dis. Chest*, 14: 633-665, 1948.

HARRIS, E. K. Unpublished dissertation. Yale University, School of Public Health, New Haven, Connecticut, 1950.

LIEBOW, A. A., G. E. LINDSKOG and W. E. BLOOMER. Cytological studies of sputum and bronchial secretions in the diagnosis of cancer of the lung. *Cancer*, 1: 223-233, 1948.

MATHER, K. Statistical Analysis in Biology. *New York*: Interscience Publishers, Inc., 1947. 267 pp.

NEYMAN, J. Contribution to the Theory of the χ^2 Test. *Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability*, Jerzy Neyman, editor. *Berkeley, California*: University of California Press, 1945-46. Pp. 239-273.

———. Outline of statistical treatment of the problem of diagnosis. *Pub. Health Rep.*, 62: 1449-1456, Oct. 3, 1947.

PAPANICOLAOU, G. N. The sexual cycle in the human female as revealed by vaginal smears. *Am. J. Anat.*, 52 (Suppl.): 533-537, 1933.

———. A new procedure for staining vaginal smears. *Science*, 95: 438-439, April 24, 1942.

USPENSKY, J. V. Introduction to Mathematical Probability. *New York*: McGraw-Hill Book Company, Inc., 1937. 411 pp.

WANDALL, H. H. A study of neoplastic cells in sputum as a contribution to the diagnosis of primary lung cancer. *Acta chir. Skandinav.* (Suppl. 93), 91: 1-143, 1944.

WOOLNER, L. B. and J. R. MCDONALD. Bronchogenic carcinoma: Diagnosis by

microscopic examination of sputum and bronchial secretions; Preliminary report. *Proc. Mayo Clinic*, 22: 369-381, Sept. 3, 1947.

———, and ———. Carcinoma cells in sputum and bronchial secretions. *Surg., Gynec. and Obst.*, 88: 273-290, 1949.

YERUSHALMY, J. Statistical problems in assessing methods of medical diagnosis, with special reference to X-ray techniques. *Pub. Health Rep.*, 62: 1432-1449, Oct. 3, 1947.

TECHNICAL APPENDIX

Derivation of equation (9): $F_t(P_k) = 0$.

Assuming equations (7), for $n = 4$, introduce the following changes of variables:

$$A_i = \frac{n!}{i!(n-i)!} P_i, \qquad (i = 0, 1, \cdots, 4), \tag{A. 1}$$

$$\alpha\pi^4 = u, \qquad (1-\alpha)(1-p)^4 = v, \tag{A. 2}$$

$$\frac{1-\pi}{\pi} = s, \qquad \frac{p}{1-p} = t.$$

Then,

$$\begin{aligned} A_0 &= u + v \\ A_1 &= su + tv \\ A_2 &= s^2u + t^2v \\ A_3 &= s^3u + t^3v \\ A_4 &= s^4u + t^4v. \end{aligned} \tag{A. 3}$$

The first three linear equations in two unknowns, u and v, can hold simultaneously only if the determinant

$$\begin{vmatrix} A_0 & 1 & 1 \\ A_1 & s & t \\ A_2 & s^2 & t^2 \end{vmatrix} = 0 \tag{A. 4}$$

for unless this condition holds in a system of $(n+1)$ equations in n unknowns, the equations are inconsistent.

Likewise, the following conditions hold:

$$\begin{vmatrix} A_1 & s & t \\ A_2 & s^2 & t^2 \\ A_3 & s^3 & t^3 \end{vmatrix} = st \begin{vmatrix} A_1 & 1 & 1 \\ A_2 & s & t \\ A_3 & s^2 & t^2 \end{vmatrix} = 0, \qquad \text{(A. 5)}$$

$$\begin{vmatrix} A_2 & s^2 & t^2 \\ A_3 & s^3 & t^3 \\ A_4 & s^4 & t^4 \end{vmatrix} = s^2t^2 \begin{vmatrix} A_2 & 1 & 1 \\ A_3 & s & t \\ A_4 & s^2 & t^2 \end{vmatrix} = 0. \qquad \text{(A. 6)}$$

These determinants are expanded to obtain the equations

$$A_2 - xA_1 + yA_0 = 0 \qquad \text{(A. 4')}$$

$$A_3 - xA_2 + yA_1 = 0 \qquad \text{(A. 5')}$$

$$A_4 - xA_3 + yA_2 = 0 \qquad \text{(A. 6')}$$

where $x = s + t$, $y = st$.

Again, to be consistent, these three equations in two unknowns, x and y, can hold simultaneously only if

$$\begin{vmatrix} A_2 & A_1 & A_0 \\ A_3 & A_2 & A_1 \\ A_4 & A_3 & A_2 \end{vmatrix} = 0. \qquad \text{(A. 7)}$$

Expanding (A. 7),

$$A_2^3 - 2A_1A_2A_3 + A_1^2A_4 + A_0A_3^2 - A_0A_2A_4 = 0. \qquad \text{(A. 8)}$$

Substituting in (A. 8) according to equations (A. 1), equation (9) is obtained.

Equations used in the estimation of the parameters, π and p, are obtained as follows:

Rewriting equations (A. 4') and (A. 5') in terms of s and t,

$$A_2 - A_1s - A_1t + A_0st = 0,$$

$$A_3 - A_2s - A_2t + A_1st = 0.$$

Considering only one variable, t for example, these two linear equations will hold simultaneously only if

$$\begin{vmatrix} (A_2 - A_1s) & (A_0s - A_1) \\ (A_3 - A_2s) & (A_1s - A_2) \end{vmatrix} = 0. \qquad \text{(A. 9)}$$

Expanding (A. 9):

$$s^2(A_0A_2 - A_1^2) - s(A_0A_3 - A_1A_2) + (A_1A_3 - A_2^2) = 0. \quad \text{(A. 10)}$$

The same result will hold in terms of t if (A. 4′) and (A. 5′) are considered with respect to the single variable s.

Dividing equation (A. 10) by $(A_0A_2 - A_1^2)$ and substituting for the A's according to equations (A. 1), one obtains the coefficients, A and B, given in the text.

The formula for the parameter α is derived from the first two equations of the system of equations (A. 3). Substitute for u, v, s, t according to equations (A. 2) and solve the resulting expression for α.